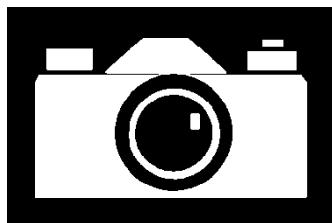


**FOCUSING ON STUDENT OUTCOMES: A
LONGITUDINAL STUDY OF DELAWARE'S
STUDENT ACCOUNTABILITY PLAN**



JANUARY 2002

LISA A. BANICKY, PH.D.
EDUCATIONAL RESEARCHER FOR POLICY ANALYSIS



Delaware Education Research & Development Center
University of Delaware
Newark, DE 19716

PUBLICATION T02.002.1
Copyright © 2002 by the University of Delaware

ACKNOWLEDGEMENTS

The author extends a special thanks to the following individuals who provided data and assistance necessary to complete this study:

Terry Anderson

Ann Case

Shirley Dear

Katie Semmel

Tommy Tao

Robin Taylor

This research was made possible through the support of the Delaware State Board of Education.

Copyright © 2002 by the University of Delaware

TABLE OF CONTENTS

List of Graphs
 List of Figures
 Executive Summary..... 4
 Introduction..... 7
 Study Design..... 7
 Report Layout..... 7
 Section 1: Special Education Placement Rates..... 9
 Section 2: Student Achievement 12
 Cross-sectional Examination of Student Test Scores 12
 Results Overview 13
 3rd Grade Summary 15
 5th Grade Summary 15
 8th Grade Summary 15
 10th Grade Summary 16
 Test score gains 28
 Performance Level Changes Between Test Administrations 31
 Section 3: Behavioral Effects 46
 Elementary Cohort Suspensions: Student Demographics 46
 Secondary Cohort Suspensions: Student Demographics 48
 Suspension and DSTP Performance 50
 Suspension History for the Cohorts 52
 Section 4: Retention 54
 Elementary Cohort: Demographic Characteristics of Retainees 54
 Secondary Cohort: Demographic Characteristics of Retainees 56
 Retention and DSTP Performance 58
 Elementary Cohort 58
 Secondary Cohort 62
 Section 5: Completion Rates..... 65
 Demographic Characteristics of Drop-outs 65
 Dropping Out and DSTP Performance..... 67
 Policy Considerations..... 70
 References..... 73
 Appendix A: Cohort Profiles..... 74
 Appendix B: Technical Details Concerning the DSTP and Statistical Interpretations..... 75
 Appendix C: Additional Graphs of Performance Levels for Several Student Groups..... 77

LIST OF GRAPHS

- Graph 1: Regular and Special Education Re-classification Rates for the Elementary Cohort
- Graph 2: Regular and Special Education Re-classification Rates for the Secondary Cohort
- Graph 3: An Examination of Re-classification Rates for 3rd Graders
- Graph 4: Special Education Categories of Re-classified Students
- Graph 5: Grade 3 Reading Standards-Based Scores by Gender
- Graph 6: Grade 3 Reading Standards-Based Scores by Race
- Graph 7: Grade 3 Reading Standards-Based Scores by Educational Service Categories
- Graph 8: Grade 3 Math Standards-Based Scores by Gender
- Graph 9: Grade 3 Math Standards-Based Scores by Race
- Graph 10: Grade 3 Math Standards-Based Scores by Educational Service Categories
- Graph 11: Grade 5 Reading Standards-Based Scores by Gender
- Graph 12: Grade 5 Reading Standards-Based Scores by Race
- Graph 13: Grade 5 Reading Standards-Based Scores by Educational Service Categories
- Graph 14: Grade 5 Math Standards-Based Scores by Gender
- Graph 15: Grade 5 Math Standards-Based Scores by Race
- Graph 16: Grade 5 Math Standards-Based Scores by Educational Service Categories
- Graph 17: Grade 8 Reading Standards-Based Scores by Gender
- Graph 18: Grade 8 Reading Standards-Based Scores by Race
- Graph 19: Grade 8 Reading Standards-Based Scores by Educational Service Categories
- Graph 20: Grade 8 Math Standards-Based Scores by Gender
- Graph 21: Grade 8 Math Standards-Based Scores by Race
- Graph 22: Grade 8 Math Standards-Based Scores by Educational Service Categories
- Graph 23: Grade 10 Reading Standards-Based Scores by Gender
- Graph 24: Grade 10 Reading Standards-Based Scores by Race
- Graph 25: Grade 10 Reading Standards-Based Scores by Educational Service Categories
- Graph 26: Grade 10 Math Standards-Based Scores by Gender
- Graph 27: Grade 10 Math Standards-Based Scores by Race
- Graph 28: Grade 10 Math Standards-Based Scores by Educational Service Categories
- Graph 29: 3rd Grade Achievement Gaps
- Graph 30: Grade 5 Achievement Gaps
- Graph 31: Grade 8 Achievement Gaps
- Graph 32: Grade 10 Achievement Gaps

- Graph 33: 3rd Grade American Indian Students
- Graph 34: 3rd Grade African American Students
- Graph 35: 3rd Grade Asian Students
- Graph 36: 3rd Grade Hispanic Students
- Graph 37: 3rd Grade Caucasian Students
- Graph 38: 3rd Grade Low Income Students
- Graph 39: 5th Grade American Indian Students
- Graph 40: 5th Grade African American Students
- Graph 41: 5th Grade Asian Students
- Graph 42: 5th Grade Hispanic Students
- Graph 43: 5th Grade Caucasian Students
- Graph 44: 5th Grade Low Income Students
- Graph 45: 8th Grade American Indian Students
- Graph 46: 8th Grade African American Students
- Graph 47: 8th Grade Asian Students
- Graph 48: 8th Grade Hispanic Students
- Graph 49: 8th Grade Caucasian Students
- Graph 50: 8th Grade Low Income Students
- Graph 51: 10th Grade American Indian Students
- Graph 52: 10th Grade African American Students
- Graph 53: 10th Grade Asian Students
- Graph 54: 10th Grade Hispanic Students
- Graph 55: 10th Grade Caucasian Students
- Graph 56: 10th Grade Low Income Students
- Graph 57: Annual Number of Out-of-School Suspensions for the Elementary Cohort
- Graph 58: Elementary Cohort: Ethnicity of Students Receiving OSS in 1999-00
- Graph 59: Elementary Cohort: Ethnicity of Students Receiving OSS in 2000-01
- Graph 60: Annual Number of Out-of-School Suspensions for the Secondary Cohort
- Graph 61: Secondary Cohort: Ethnicity of Students Receiving OSS in 1999-00
- Graph 62: Secondary Cohort: Ethnicity of Students Receiving OSS in 2000-01
- Graph 63: Elementary Cohort: Spring 2000 DSTP Performance for Students Receiving OSS in 1999-00

Graph 64: Secondary Cohort: Spring 2000 DSTP Performance for Students Receiving OSS in 1999-00

Graph 65: Suspension History for Students Suspended During the 2000-01 School Year (Elementary Cohort)

Graph 66: Suspension History for Students Suspended During the 2000-01 School Year (Secondary Cohort)

Graph 67: Demographic Information for Students Retained in Grade 4

Graph 68: Ethnicity of Students Retained in Grade 4

Graph 69: Demographic Information for Students Retained in Grade 5

Graph 70: Ethnicity of Students Retained in Grade 5

Graph 71: Demographic Information for Students Retained in Grade 7

Graph 72: Ethnicity of Students Retained in Grade 7

Graph 73: Demographic Information for Students Retained in Grade 8

Graph 74: Ethnicity of Students Retained in Grade 8

Graph 75: 2000 DSTP Reading Performance for Retained Students (Elementary Cohort)

Graph 76: 2001 DSTP Reading Performance for Retained Students (Elementary Cohort)

Graph 77: 2000 DSTP Math Performance for Retained Students (Elementary Cohort)

Graph 78: 2001 DSTP Math Performance for Retained Students (Elementary Cohort)

Graph 79: 2000 DSTP Writing Performance for Retained Students (Elementary Cohort)

Graph 80: 2001 DSTP Writing Performance for Retained Students (Elementary Cohort)

Graph 81: 2000 DSTP Reading Performance for Retained Students (Secondary Cohort)

Graph 82: 2001 DSTP Reading Performance for Retained Students (Secondary Cohort)

Graph 83: 2000 DSTP Math Performance for Retained Students (Secondary Cohort)

Graph 84: 2001 DSTP Math Performance for Retained Students (Secondary Cohort)

Graph 85: 2000 DSTP Writing Performance for Retained Students (Secondary Cohort)

Graph 86: 2001 DSTP Writing Performance for Retained Students (Secondary Cohort)

Graph 87: Demographic Characteristics of Students Who Dropped out during 2000-01

Graph 88: Ethnicity of 8th Grade Drop-outs

Graph 89: Promotion History of Eventual Drop-outs

Graph 90: Reading Performance Levels for Eventual Drop-outs (2000-01)

Graph 91: Math Performance Levels for Eventual Drop-outs (2000-01)

Graph 92: Writing Performance Levels for Eventual Drop-outs (2000-2001)

LIST OF FIGURES

Figure 1: An Overview of 3rd Grade DSTP Results

Figure 2: An Overview of 5th Grade DSTP Results

Figure 3: An Overview of 8th Grade DSTP Results

Figure 4: An Overview of 10th Grade DSTP Results

Figure 5: Percent of Students At or Above the Standard in Reading – Spring 2001

Figure 6: Percent of Students At or Above the Standard in Math – Spring 2001

EXECUTIVE SUMMARY

Most states and local school districts across the nation have implemented high-stakes assessment systems to gather information about student achievement for the purpose of holding schools, teachers, and students accountable. The cornerstones of such systems include content standard, assessments, and consequences for performance. Many educational experts agree that the consequences included in some high stakes assessment systems can lead to unintended consequences and manipulations of the system. These experts recommend that such systems be routinely monitored for both intended and unintended consequences.

Based on the recommendations of several national educational organizations and existing research, several key issues were addressed in light of the accountability plan. This report represents the second in a series of annual reports designed to monitor the impact of Delaware's Student Accountability Plan on students in the First State. The previous years findings and the follow-up findings related to the key issues are summarized in the following tables:

SPECIAL EDUCATION PLACEMENT RATES

- **Previous Findings:** From 1997-2000 less than 4% of regular education students were later re-classified as special education students. However, large percentages of special education students were re-classified as regular education when comparing Fall to Spring enrollment information.
- **Follow-up Findings:** When comparing Spring 2000 to Fall 2001 special education placements, less than 5% of regular education students were re-classified as special education students and less than 5% of special education students were re-classified as regular education. The large percentages of re-classifications for special education students reported last year appear to have resulted from testing policies in which students with speech impairments were re-classified as regular education for the purposes of the DSTP.

STUDENT ACHIEVEMENT

Previous Findings: Based on data from the **elementary cohort**, SAT9 reading and DSTP writing scores did not improve over time (Spring 1998 to Spring 2000) but the SAT9 mathematics scores showed slight improvements for most student groups. With respect to standards-based scores, an examination of student performance levels indicated that more students were performing at or above the standard in reading and math and fewer students were performing at or above the standard in writing in 2000 as compared to 1998.

Follow-up Findings: A cross-sectional examination of student performance on the DSTP was conducted based on the Spring 1998 and Spring 2001 administration. An examination of the standards-based scores revealed that most of the student groups at the elementary level demonstrated significant improvements in the areas of reading and mathematics. Special Education and Title I students evidenced improvements in reading and mathematics in grades 3, 5, and 8. None of the student groups at the 10th grade level showed improvement in the area of reading and only a few showed improvement in mathematics. An examination of the percentage of students at each of the performance levels revealed that gaps in achievement continue to persist for African American, Hispanic, and Low Income students. The percentage of students at or above the standard in Spring 2001 varied according to student group, grade, and content area.

BEHAVIORAL EFFECTS

- **Previous Findings:** For students in both the elementary and secondary cohorts, the odds of receiving an out of school suspension were greater for male, low income status, and minority students. Although not statistically significant, suspended students had lower DSTP scores compared to their non-suspended peers.
- **Follow-up Findings:** The number of students receiving out-of-school suspensions increased among the elementary cohort but decreased among the secondary cohort. The odds of suspension were again found to be greater for male, low income status, and minority students. In both cohorts a high number of repeat offenders were found. With respect to DSTP performance, a majority of students receiving suspensions during the 1999-00 school year performed below state content standards in reading and mathematics.

RETENTION

- **Previous Findings:** For students in both the elementary and secondary cohorts, the odds of being retained increased for certain student groups. This finding was qualified by whether a student had actually been retained or would have been retained if the consequences associated with the DSTP had been in effect. When considering potential retentions, the odds of being retained were greater for low income status, minority, and male students. The DSTP scores for both actual and potential retainees indicated that a majority of these students continued to perform below the standard when re-taking portions of the DSTP.
- **Follow-up Findings:** The odds of being retained were again found to be higher for males, minority students, and low-income students in both the elementary and secondary cohorts. An examination of the performance levels of retained students revealed increases in the percentages of students meeting the standard in the following spring (2001) administration of the DSTP. However, large percentages of the retained students were below the standard in math and more than one-third of the retained students were below the standard in reading and writing in 2001.

COMPLETION RATES

- **Current Findings:** Within the secondary cohort there were officially 4 students who dropped-out during the 1999-00 school year and 14 students who dropped out in 2000-01. Given the small number of official dropouts the relationship of demographics to dropping out and the relationship between DSTP performance and dropping out were examined descriptively. The tentative results suggest that low income students may be more likely to drop out and that students with a history of retention may be more likely to drop-out. For the few drop-out students with DSTP data available, the results indicated that most of the eventual drop-outs were performing below the standard in reading, mathematics, and writing.

INTRODUCTION

This report is the second in a series of several annual reports designed to monitor the impact of Delaware's Student Accountability Plan on students in Delaware. More than ten years of educational reform in Delaware has led to the creation of a performance based accountability system composed of rigorous standards, a statewide assessment, and soon to be included consequences for performance.

With increased accountability for performance comes the responsibility of monitoring systems for both intended and unintended consequences. According to the National Research Council: *“high stakes testing programs [should] routinely include a well-designed evaluation component. Policymakers should monitor both the intended and unintended consequences of high-stakes assessment on all students and on significant sub-groups of students including minorities, English language learners, and students with disabilities”* (p. 281).¹

This ongoing study, conducted at the request of the Delaware State Board of Education, is intended to monitor Delaware's student accountability plan and provide ongoing information about its effects on students. Based on previous research in the area of high stakes testing, several key issues were identified during the planning of the study as areas for monitoring. These issues included:

- Special Education Placement Rates;
- Student achievement;
- Behavioral Effects;
- Retention Rates; and,
- Completion Rates.

STUDY DESIGN

To address these issues, students enrolled in 3rd or 6th grade during the 1997-98 school year were selected as cohorts for study. Additional information on the students included in the two cohorts appear in Appendix A. New to this year's report was an examination of student achievement using a cross-sectional approach in which the performance of students at the first administration of the DSTP (Spring 1998) was compared to the performance of students at the most recent administration of the DSTP (Spring 2001).

Both of these approaches represent a longitudinal approach to examining student outcomes with the first focusing on the same students over time and the second examining different groups of students at two different points in time. As an analogy, the use of cohorts provides a means of producing a “documentary” of student outcomes which follows the

same group of individuals over time whereas the cross-sectional approach allows the researcher to take a “snapshot” of student outcomes using different students at different times.

REPORT LAYOUT

In order to distinguish between the results based on the cohorts and the results based on the cross-sectional groups of students, the following icon will be used to identify results based



on the cross-sectional groups. In addition, the results have been organized around the five key issues listed above. A summary of the findings for each of these issues and the policy considerations associated with the findings appear at the end of the report in the section entitled “Policy Considerations”.

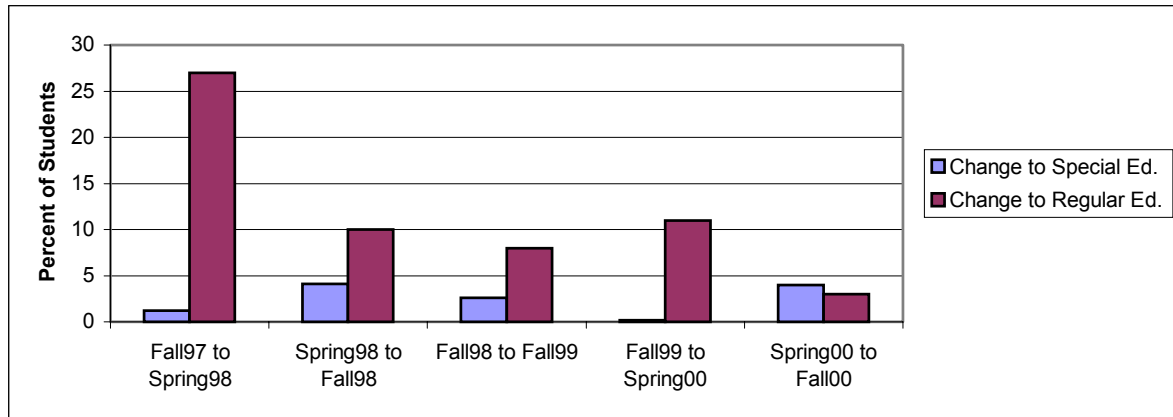
SECTION 1: SPECIAL EDUCATION PLACEMENT RATES

One unintended consequence that has been found to emerge in some high-stakes assessment systems is an increase in special education enrollment rates.² Such increases may represent an attempt to manipulate the system and exclude students from testing that are expected to perform poorly. With recent changes to testing policies and added measures to include students with disabilities such manipulations should be less likely to occur.

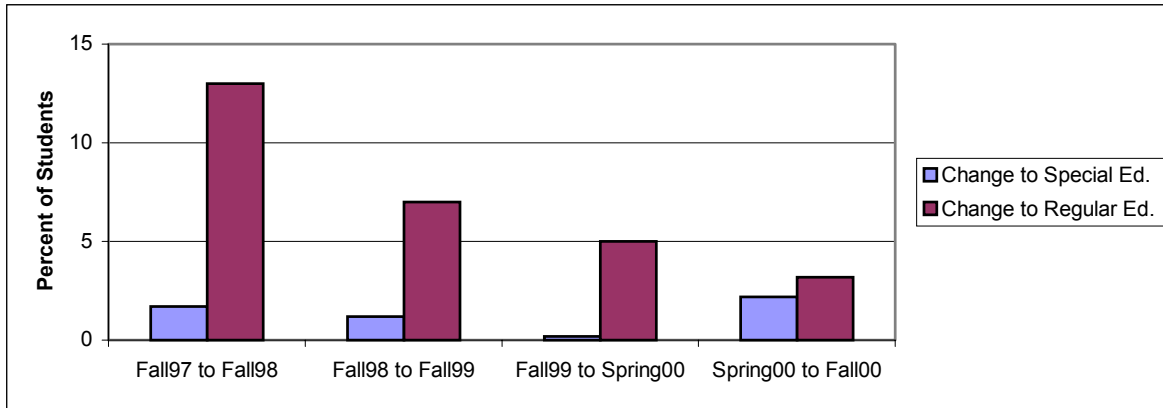
For the purposes of this study, the guiding question with respect to the issue of special education placement rates was: **Are the special education placement rates changing within the two cohorts over time, and if so in what ways?**

The re-classification rates for the elementary and secondary cohort appear in Graphs 1 and 2 below. For both the elementary and secondary cohort the largest changes in special education placement rates involved students who were re-classified from special education into regular education. The highest re-classification rates for the elementary and secondary cohort occurred from the Fall of 1997 to the Spring of 1998. Most notable was the re-classification of 27% of the special education students in the elementary cohort at the time of the Spring 1998 DSTP administration. The percentage of regular education students re-classified as special education did not exceed 5% for any of the time periods under investigation.

Graph 1: Regular and Special Education Re-classification Rates for the Elementary Cohort



Graph 2: Regular and Special Education Re-classification Rates for the Secondary Cohort



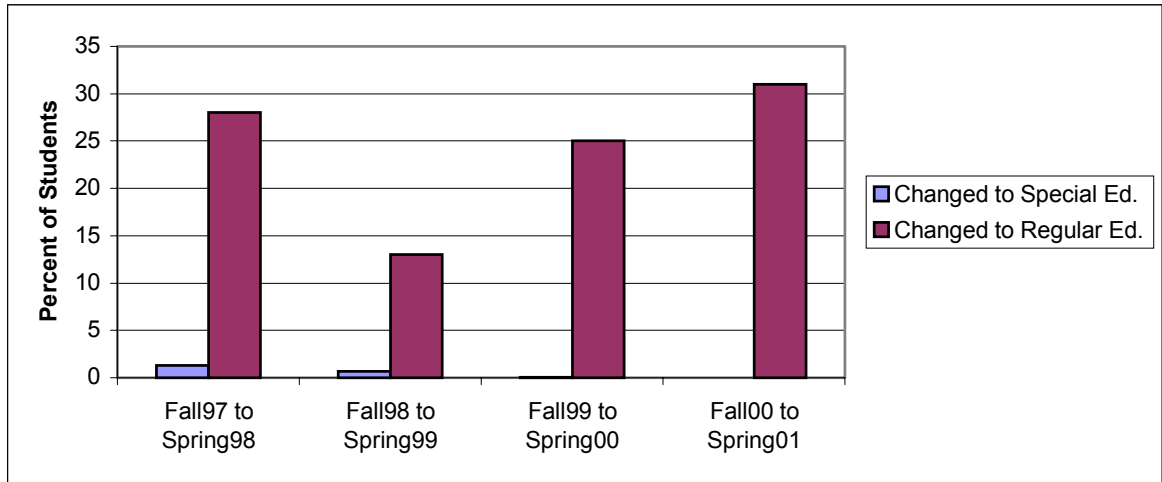
During the course of the year, with several opportunities to share the results of last year’s study with a variety of educator groups, several explanations for these findings were offered. The explanations included developmental issues specific to the early elementary grades, pressure from the Office of Civil Rights during the 1997-98 school year to reduce the number of special education classifications, and policy decisions concerning the classification of certain disabilities for the purpose of aggregating DSTP test scores.

While the role of developmental growth could not be directly addressed with the data available, there were data available to address the other explanations for the placement rates. If the increase in special education students re-classified as regular education was the result of pressure from the Office of Civil Rights during the 1997-98 school year, then the high percentage of re-classifications should be isolated to that particular year. If the number of special education students re-classified represented a testing policy applied each time the test was administered then the pattern should reappear in subsequent testing years. According to several members of the Delaware Department of Education a decision had been made to classify students with speech impairments as regular education for DSTP testing purposes.³

Cross-sectional data of 3rd graders from 1997-2000 were examined to determine if the high re-classification rate for special education students was specific to the 1997-98 school year or if the pattern occurred in other years as well. The results of this examination appear in Graph 3.

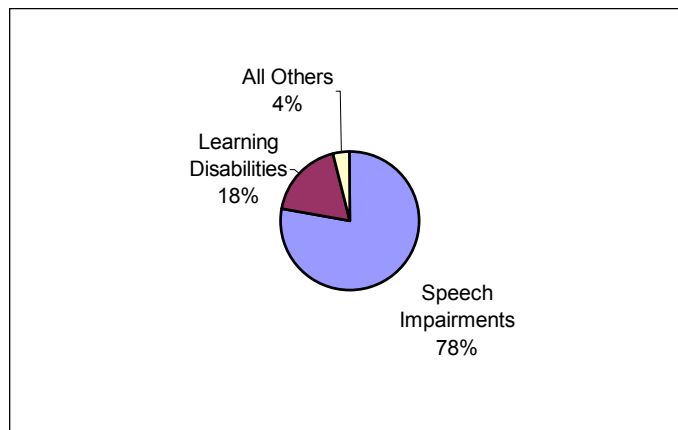


Graph 3: An Examination of Re-classification Rates for 3rd Graders



Graph 3 indicates that the large percentage of special education students re-classified as regular education was not specific to the 1997-1998 school year but represents a pattern of re-classifications. In order to determine if the students whose classification changed were predominantly speech impaired students, an examination of the type of classification (learning disability, hearing impaired, etc.) was conducted (See Graph 4).

Graph 4: Special Education Categories of Re-classified Students



The results indicated that more than two-thirds of the students re-classified during the 1997-98 school year were students with speech impairments. Similar results were found for each of the subsequent years represented in Graph 3. In addition, during the 1998-99 school year in which the re-classification of special education students to regular education was lowest, there were correspondingly lower percentages of special education students classified with a speech impairment at the beginning of the school year. Taken together these results support the testing policy explanation for the higher percentage of re-classifications during the testing year.

SECTION 2: STUDENT ACHIEVEMENT



According to Delaware Policymakers, one of the overriding goals of the Student Accountability System was to improve student achievement by providing a system for measuring student performance against state content standards.⁴ The measurement tool in this case is the Delaware Student Testing Program (DSTP) which provides results in the form of national percentile ranks, standards-based scores, and performance levels.

As of the 2000-01 school year, students in the cohorts who have progressed through the system on time should have been enrolled in grades 6 and 9, respectively. For this reason, a majority of the students in the cohort did not take the DSTP during the Spring 2001 administration. Therefore, a longitudinal examination of student achievement was conducted using a cross-sectional approach comparing students taking the DSTP during the first year of the test administration (Spring 1998) to students taking the DSTP at the most recent administration (Spring 2001). The results related to student achievement are organized around three guiding questions. These questions included:

- **Are students' test scores improving over time?**
- **Do test score gains differ as a function of gender, race, and income status?**
- **How much movement is there in performance levels on the DSTP from one test administration to the next?**

CROSS-SECTIONAL EXAMINATION OF STUDENT TEST SCORES

An examination of improvements over time was conducted through the use of independent samples t-tests. The analyses involved data disaggregated according to race, special education status, low income status, Title I status, LEP status, and gender. Student performance was examined in terms of standards-based scores in reading, mathematics, and writing. SAT9 NCEs for reading and mathematics were also examined. Students enrolled as of September 30th of the tested year were included in the analyses.

Because of the large size of several of the student groups and the large number of comparisons made on the same set of data, appropriate adjustments were made when selecting the statistical level of significance. In addition, effect sizes were calculated to determine if a statistically significant finding represented a meaningful difference. Several technical details concerning the DSTP as well as caveats concerning statistical interpretations were highlighted in last year's report and will not be repeated here. Relevant excerpts appear in Appendix B.

Although the analyses of student achievement involved the standards-based scores and SAT9 NCEs for three content areas (reading, mathematics, and writing), the decision was made to only provide detailed information concerning student performance with regard to reading and mathematics standards-based scores. This decision was based on several factors:

1. Changes were made in the scoring of the DSTP writing assessment between the first and most recent test administration which would confound the statistical interpretations.⁵
2. The goal of the accountability system is to measure student performance against state content standards. The SAT9 results in the areas of reading and mathematics provide a measure of student performance against national norms and does not directly address the attainment of state content standards.

Results Overview

The following figures were designed to provide an overview of the results of the independent t-tests. Highlighted boxes indicate areas in which a statistically significant change was found in student test scores between the Spring of 1998 and the Spring of 2001. In addition to meeting the statistical level of significance, the size of the effect had to be greater than .20 in order for the box to be highlighted. The lighter shading represents a statistically significant *improvement* while the darker shading represents a statistically significant *decline* in scores.

Figure 1: An Overview of 3rd Grade DSTP Results

	RSBS	RSAT9	MSBS	MSAT9	Writing
Male					
Female					
American Indian					
African American					
Asian					
Hispanic					
Caucasian					
Special Ed.					
LEP					
Title I					
Low Income					

*Light shading indicates improvement, dark shading decline, and no shading indicates no change.

Figure 2: An Overview of 5th Grade DSTP Results

	RSBS	RSAT9	MSBS	MSAT9	Writing
Male					
Female					
American Indian					
African American					
Asian					
Hispanic					
Caucasian					
Special Ed.					
LEP					
Title I					
Low Income					

Figure 3: An Overview of 8th Grade DSTP Results

	RSBS	RSAT9	MSBS	MSAT9	Writing
Male					
Female					
American Indian					
African American					
Asian					
Hispanic					
Caucasian					
Special Ed.					
LEP					
Title I					
Low Income					

Figure 4: An Overview of 10th Grade DSTP Results

	RSBS	RSAT9	MSBS	MSAT9	Writing
Male					
Female					
American Indian					
African American					
Asian					
Hispanic					
Caucasian					
Special Ed.					
LEP					
Title I					
Low Income					

3rd Grade Summary

With respect to the results for the 3rd grade students, the analyses revealed statistically significant increases in the area of reading and mathematics for nearly all student groups excluding American Indians.* Significant declines in the area of writing were found for nearly all student groups except for American Indians and LEP students who constitute the two smallest student groups.

5th Grade Summary

The results of the analyses based on 5th grade students again revealed statistically significant improvements in reading and mathematics for most of the student groups except for American Indians, Asians, and LEP students who showed no real change between the Spring 1998 and Spring 2001 administration. The analyses also revealed that the writing scores were not significantly different between the two test administrations.

8th Grade Summary

The analyses of the 8th grade DSTP performance indicated that only a few of the student groups showed improvement in reading and mathematics. Notable exceptions to this

* The size of the American Indian subgroup is very small which may affect the likelihood of detecting a significant effect. This is true for each of the grades under consideration in the current analyses.

finding were Hispanic students in the areas of reading and mathematics and African Americans and Low income students in the area of mathematics. These historically underperforming students posted improvements in these areas. Special education students and Title I students also showed significant improvements in all three of the content areas under study.

10th Grade Summary

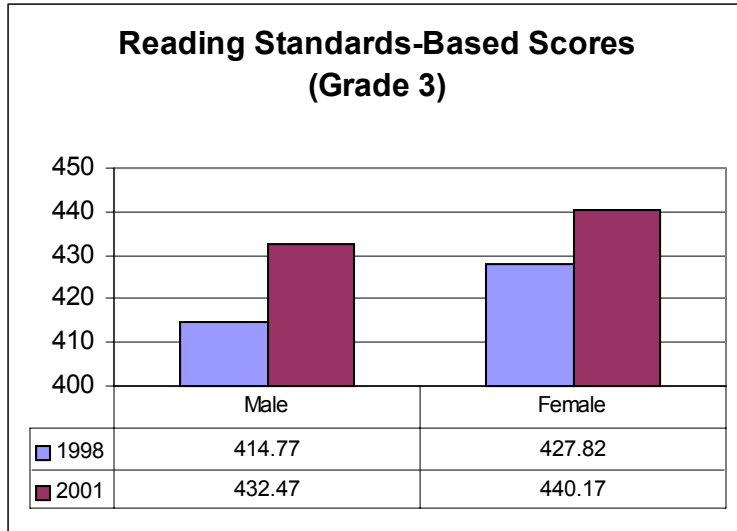
An examination of the 10th grade DSTP results revealed that none of the student groups evidenced improvements in the area of reading and only a few student groups showed improvements in the area of mathematics (i.e. African Americans, Caucasians, and Special Education Students). In addition, several of the student groups showed improvements in the areas of writing. These groups included African Americans, Asians, Caucasians, and Title I students.

Overall, in response to the original question posed, “are students’ test scores improving over time” the answer appears to be yes for students in the elementary grades but less so for students in the secondary grades. More of the students groups in the early grades showed improvement in the areas of reading and mathematics, but most of the student groups did not fare as well at the secondary level. The findings related to the area of writing were mixed with declines in some grades (3rd grade) and increases in others (10th grade).

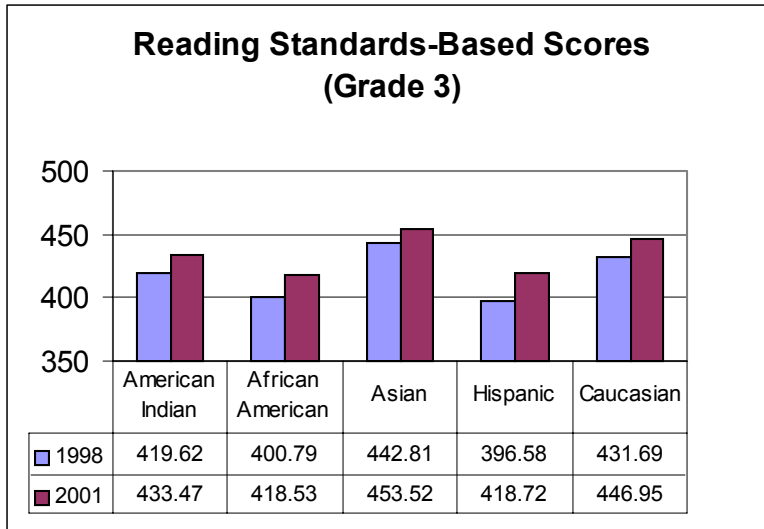
Another pattern to emerge from the data was the finding that special education and Title I students in grades 3, 5, and 8 showed significant improvements in reading and mathematics. The findings for the Title I students were associated with some of the largest effect sizes.

Graphs depicting the average reading and mathematics standards-based score for each student subgroup are presented on the following pages.

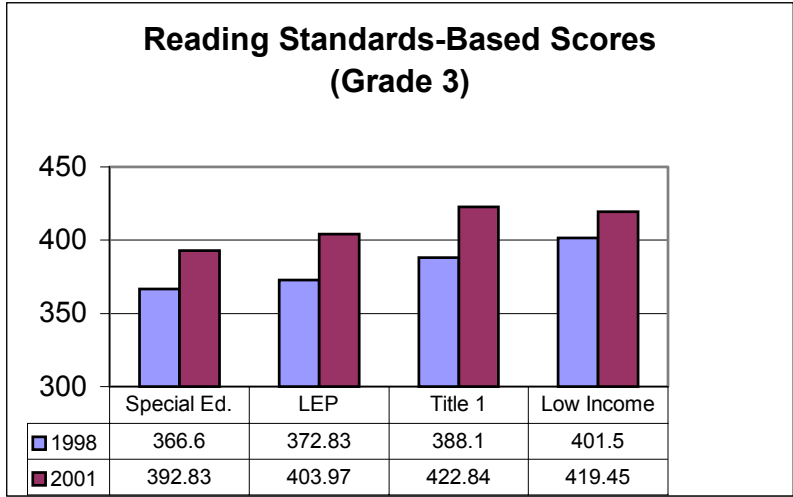
Graph 5: Grade 3—Gender



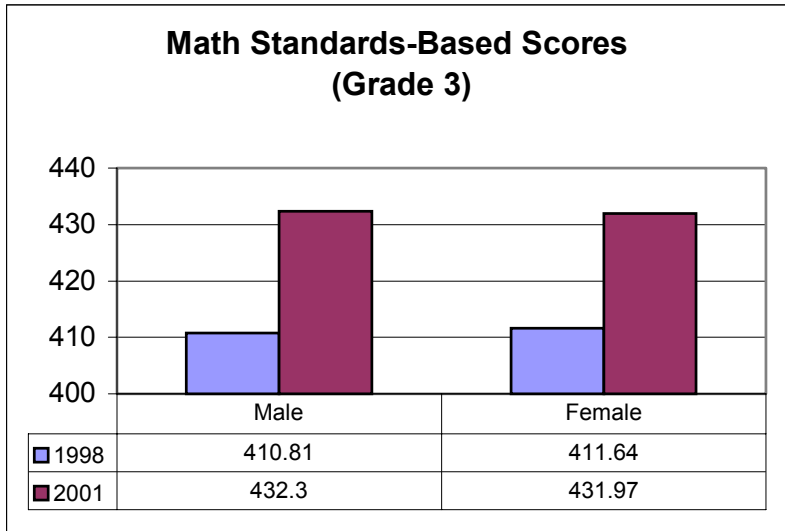
Graph 6: Grade 3--Race



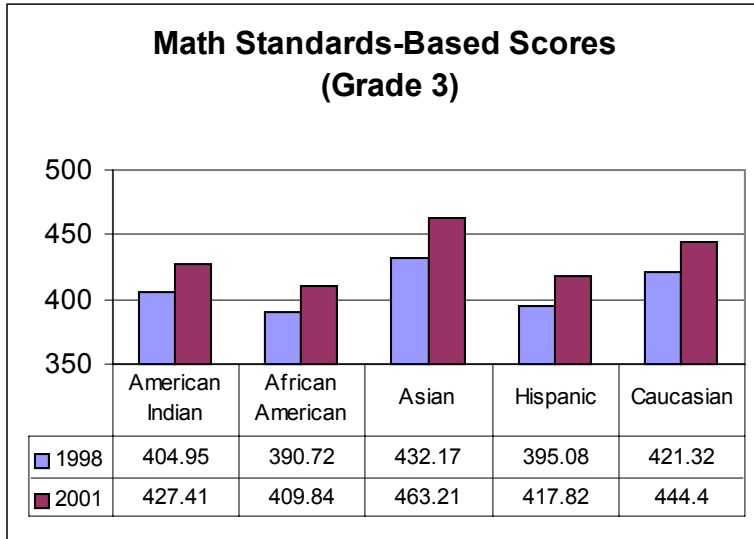
Graph 7: Grade 3--Educational Service Categories



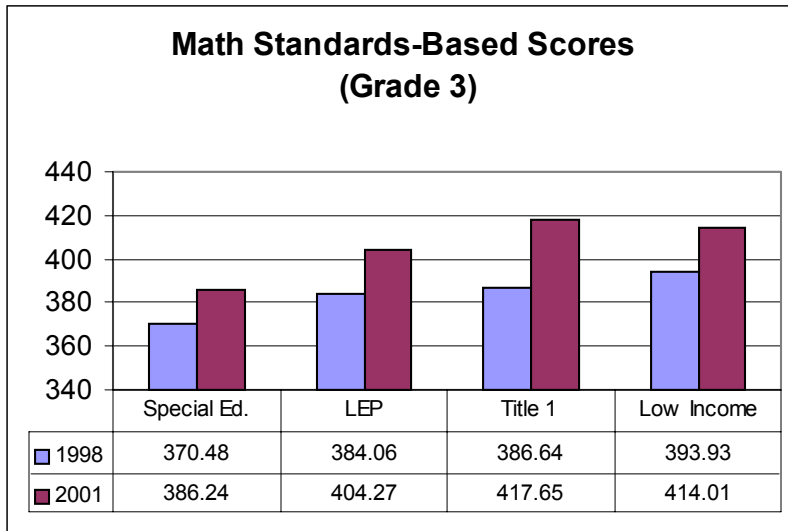
Graph 8: Grade 3—Gender



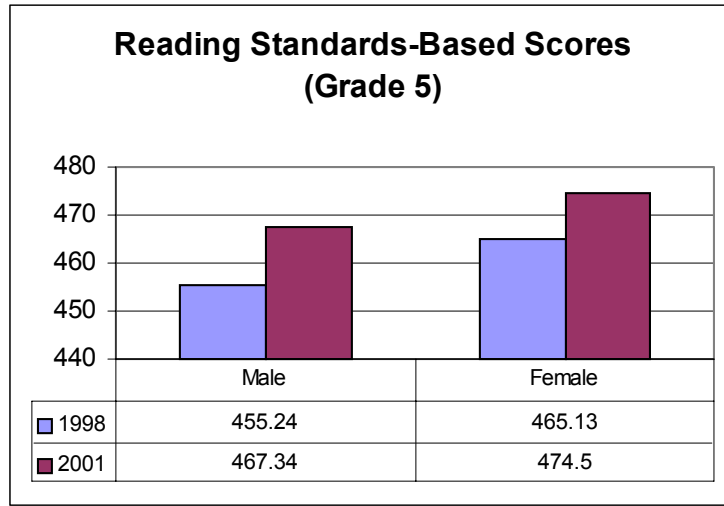
Graph 9: Grade 3—Race



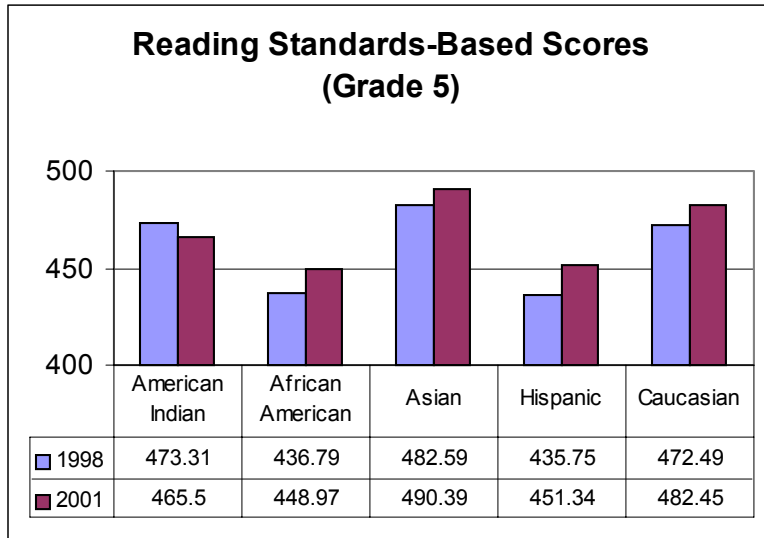
Graph 10: Grade 3—Education Service Categories



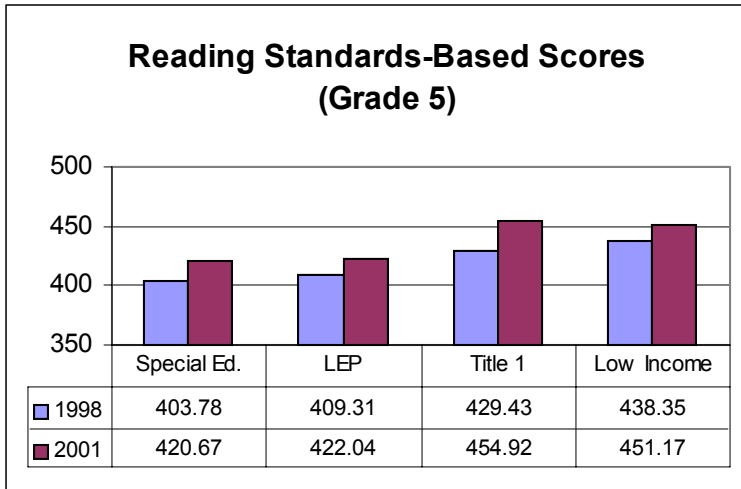
Graph 11: Grade 5—Gender



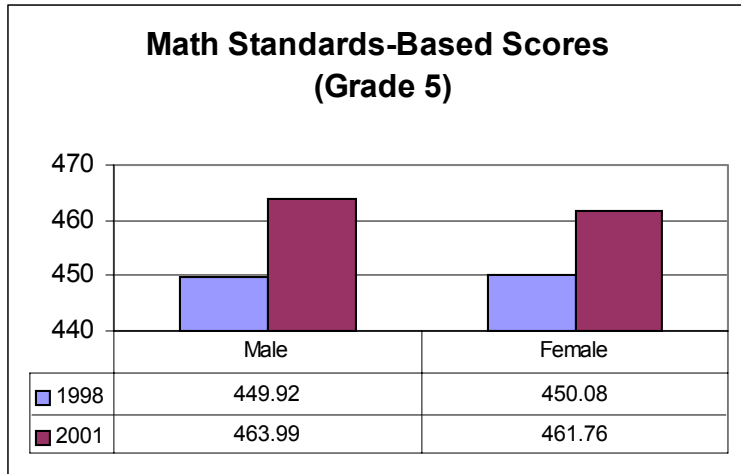
Graph 12: Grade 5—Race



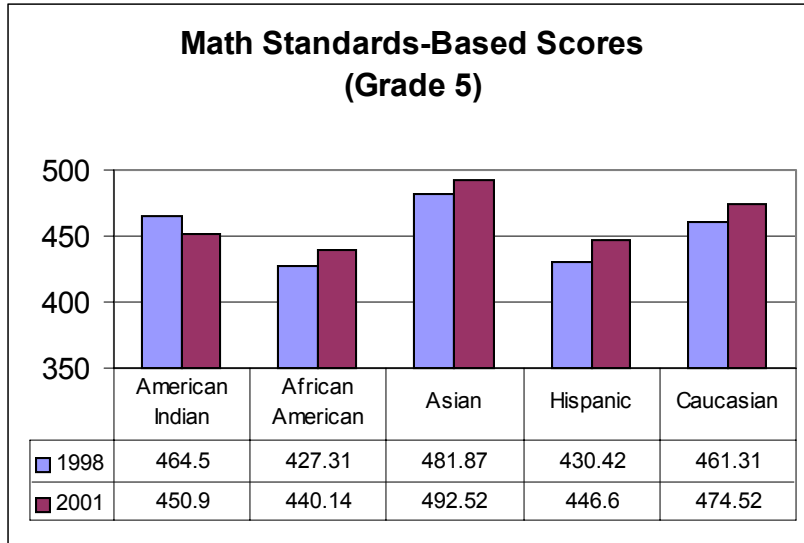
Graph 13: Grade 5—Education Service Category



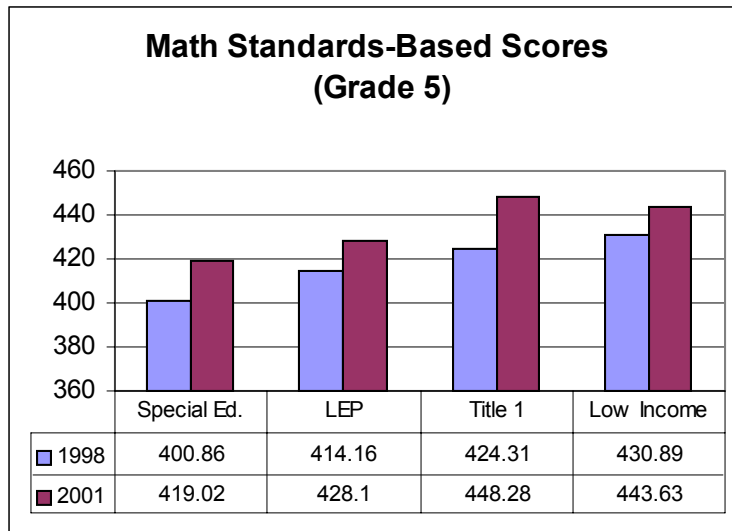
Graph 14: Grade 5—Gender



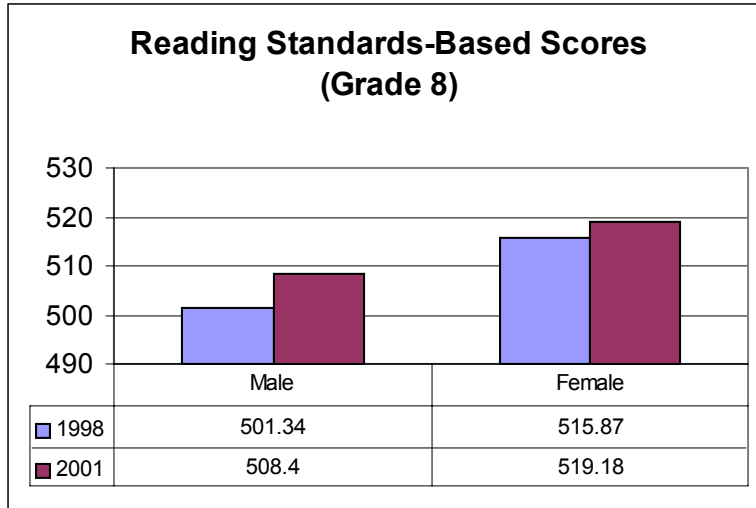
Graph 15: Grade 5—Race



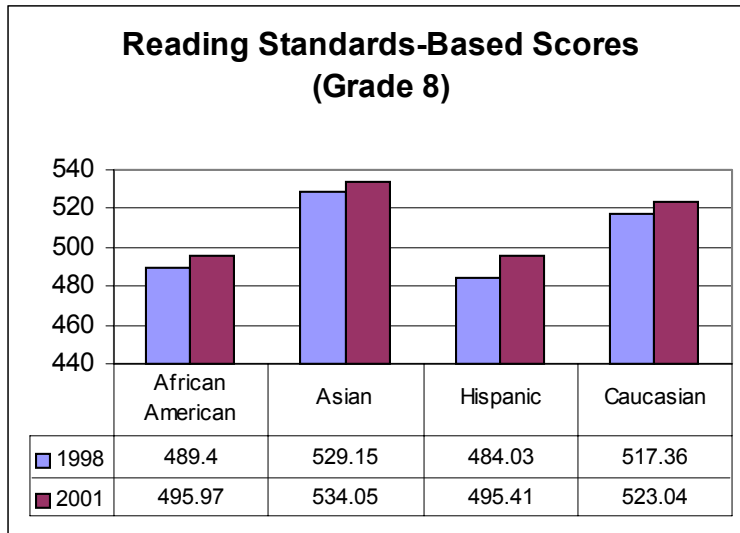
Graph 16: Grade 5—Education Service Categories



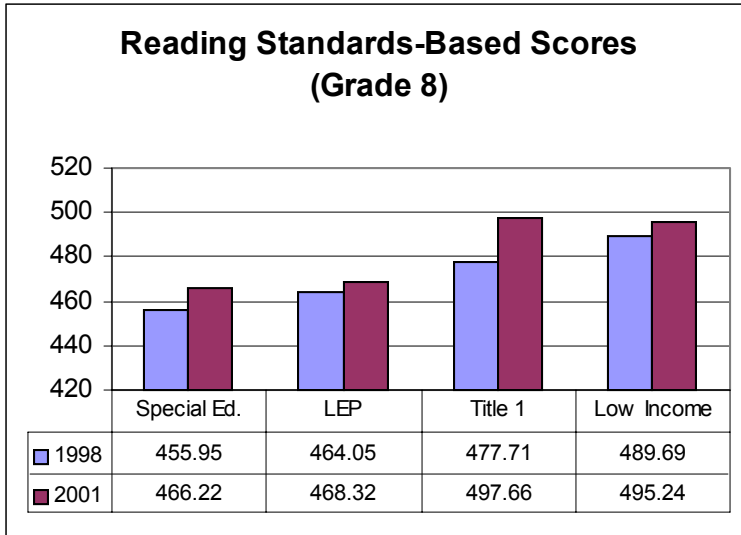
Graph 17: Grade 8—Gender



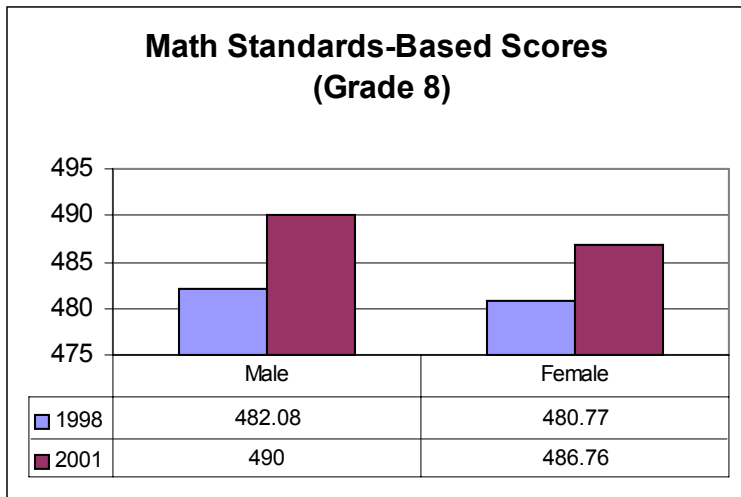
Graph 18: Grade 8—Race



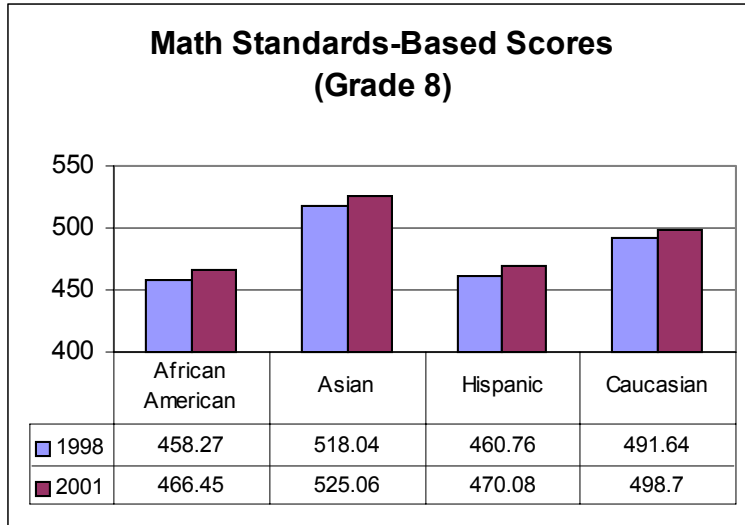
Graph 19: Grade 8—Education Service Categories



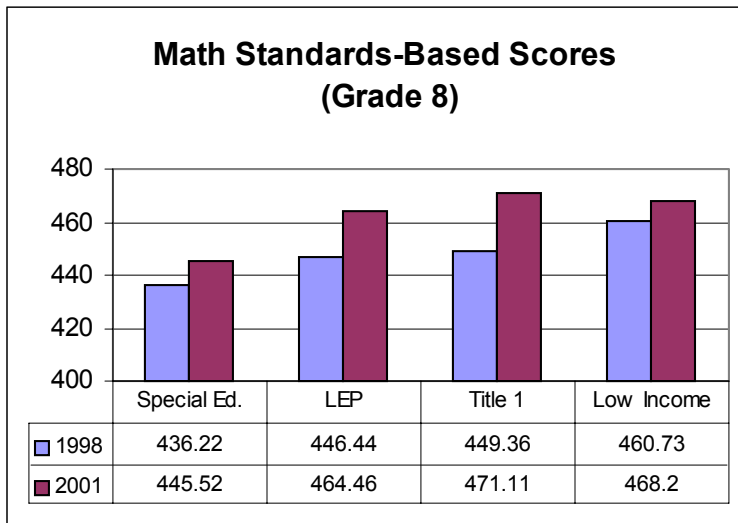
Graph 20: Grade 8—Gender



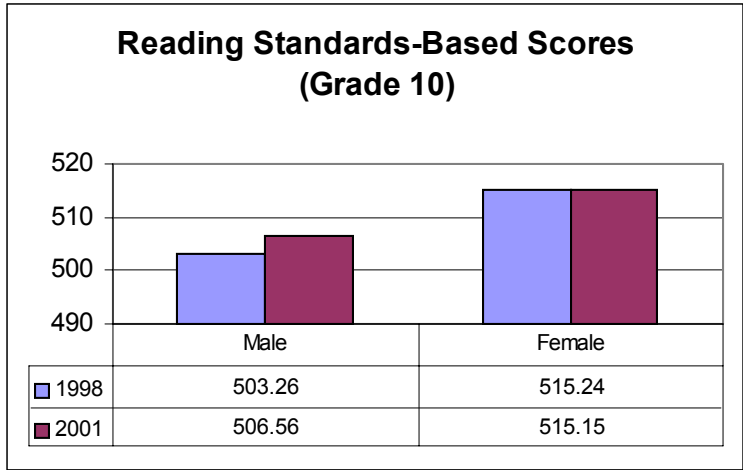
Graph 21: Grade 8—Race



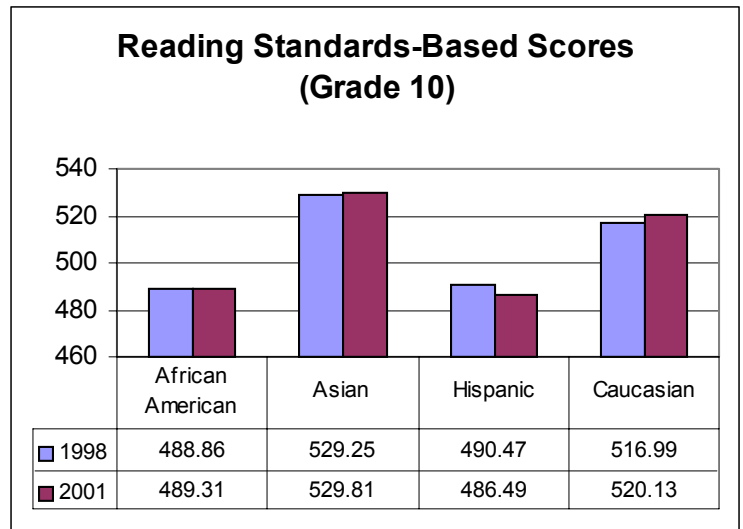
Graph 22: Grade 8—Education Service Categories



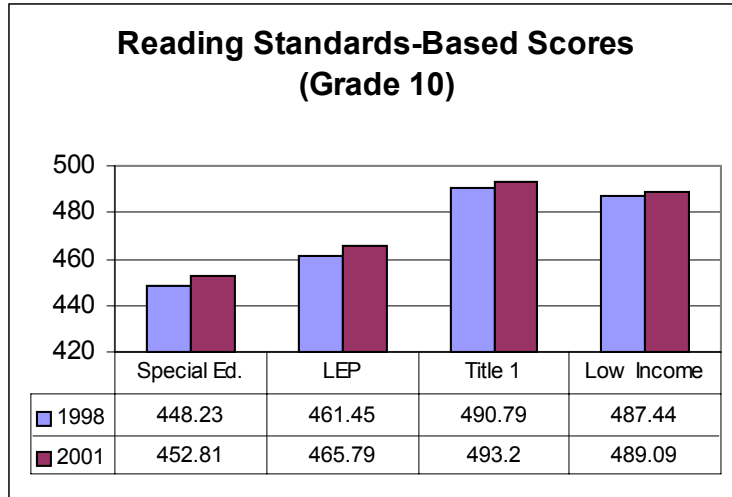
Graph 23: Grade 10—Gender



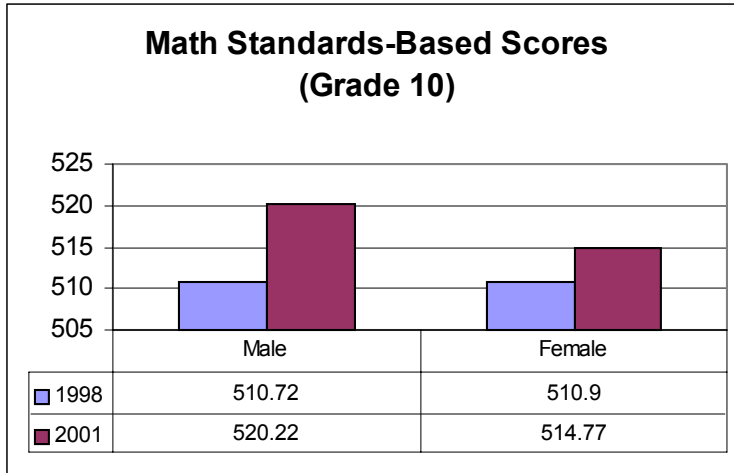
Graph 24: Grade 10—Race



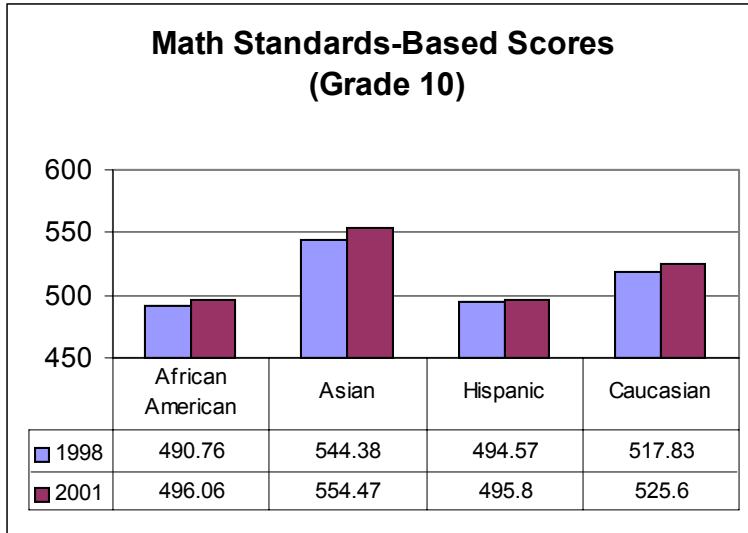
Graph 25: Grade 10—Education Service Categories



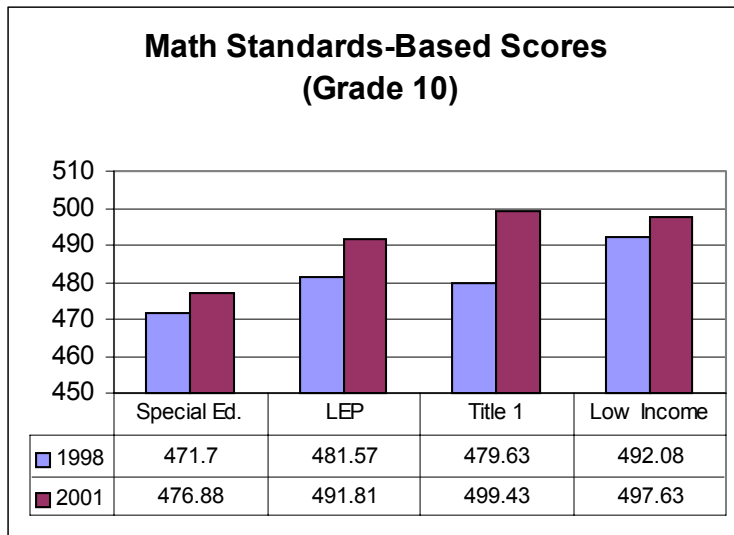
Graph 26: Grade 10—Gender



Graph 27: Grade 10—Race



Graph 28: Grade 10—Education Service Categories



TEST SCORE GAINS

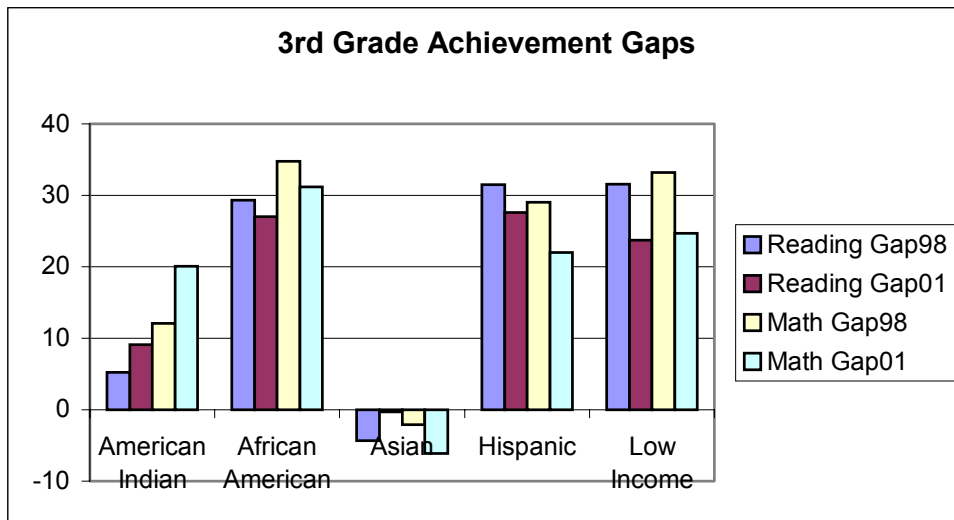
In addition to determining if student test scores were improving over time, last year’s longitudinal report examined student test score gains as a function of gender, race, and income status. Such an analysis requires the computation of a change score which is the difference between a student’s performance in one year and the same student’s performance in a subsequent testing year. The cross-sectional data were not amenable to such analyses

but cursory analyses of the gap in achievement between minorities and non-minorities and the gap between low income and non-low income students were conducted.*

The descriptive analyses that follow were based on a definition of the achievement gap developed by a working group initiated by the Delaware State Board of Education. The gap was examined by looking at the difference in the percentage of students at or above the standard for each minority group relative to Caucasians, and by looking at the percentage of low income students at or above the standard relative to non-low income students.

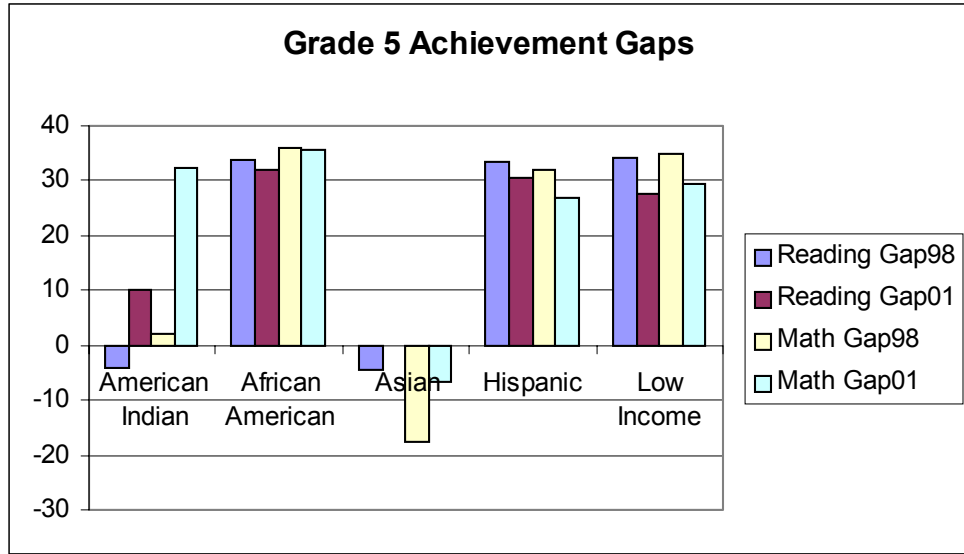
The following graphs depict the size of the achievement gaps for the areas of reading and mathematics and provide some indication of the change in the size of the gap from the Spring 1998 to the Spring of 2001. The higher the bar is above 0, the larger the discrepancy in performance between the student group listed and the appropriate comparison group. Bars that fall below the 0 line indicate a gap in favor of the minority group listed.

Graph 29

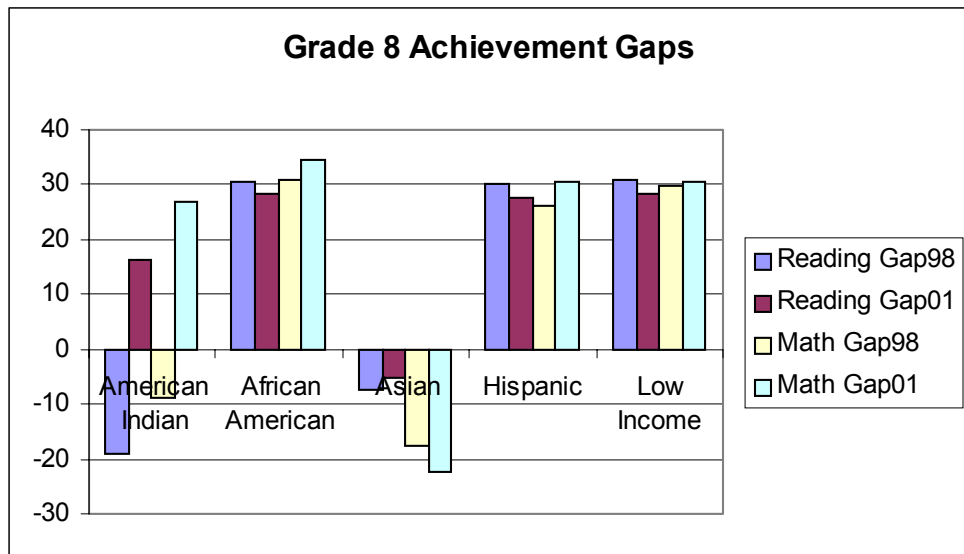


* These groups were selected based on recent state-wide initiatives to address the achievement gaps between these student groups. A more detailed analysis of achievement gaps at the individual school level is forthcoming.

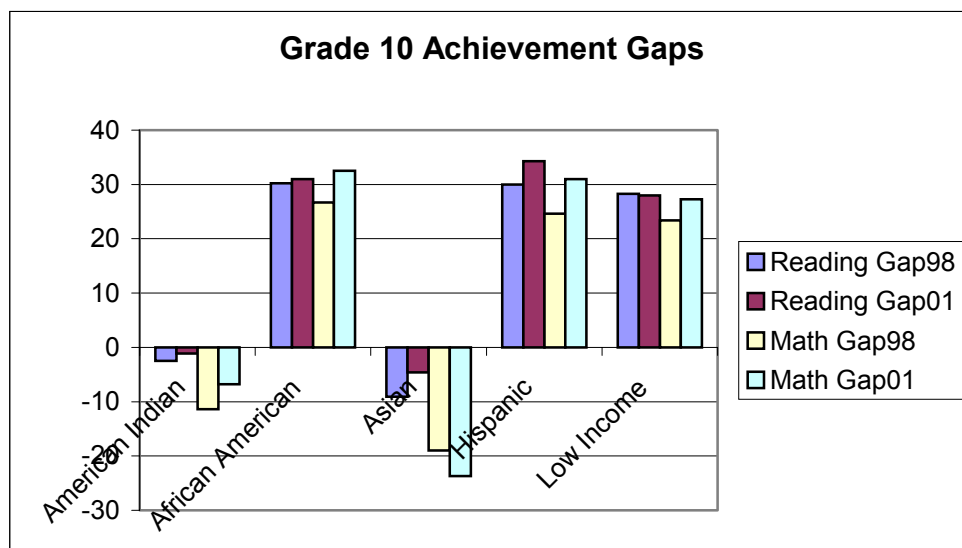
Graph 30



Graph 31



Graph 32



A cursory review of the achievement gap revealed large gaps in achievement for African American, Hispanic, and Low Income students. Across all four grade levels Asian students were the only group to outperform Caucasians. An examination of the changes to the size of the gaps revealed some decreasing gaps in reading and mathematics for African Americans, Hispanics, and Low Income students, particularly in the elementary grades. At the secondary level the gaps in achievement for these students appeared to be increasing particularly in the area of mathematics.

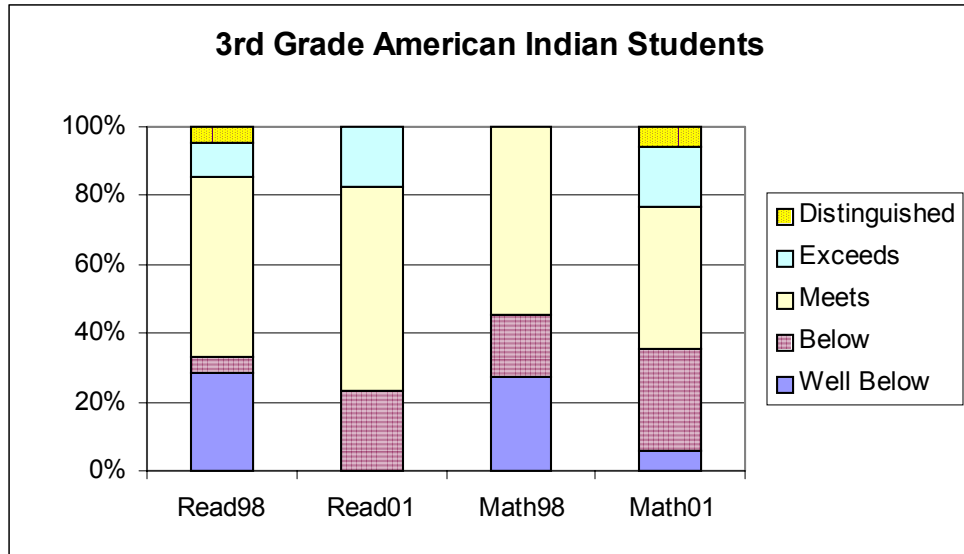
PERFORMANCE LEVEL CHANGES BETWEEN TEST ADMINISTRATIONS

In addition to looking at the size of the gap it is important to examine the absolute performance levels of the various student groups in order to determine if the narrowing gaps reflect improvement for all groups included. A gap may decrease because both student groups improve their performance but the minority group improves at a faster rate, or a gap may decrease at the expense of one student groups' performance declining over time while the minority group's performance remains unchanged. In order to have the smaller gap represent a desired change, the absolute performance of both student groups should improve over time.

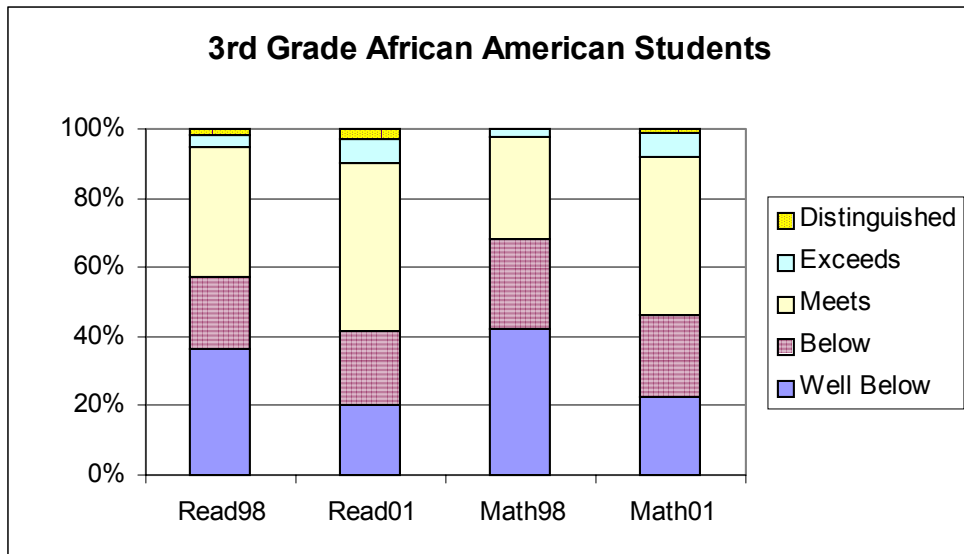
Descriptive statistics were used to examine the absolute levels of performance for all student groups in the areas of reading and mathematics. The following graphs show the percentage of students performing at each of the performance levels in the Spring of 1998 and the Spring of 2001. Although the gap indices were only computed based on racial subgroups

and income status, the performance of all student groups were examined and the graphs representing absolute performance of the other student groups appear in Appendix C.

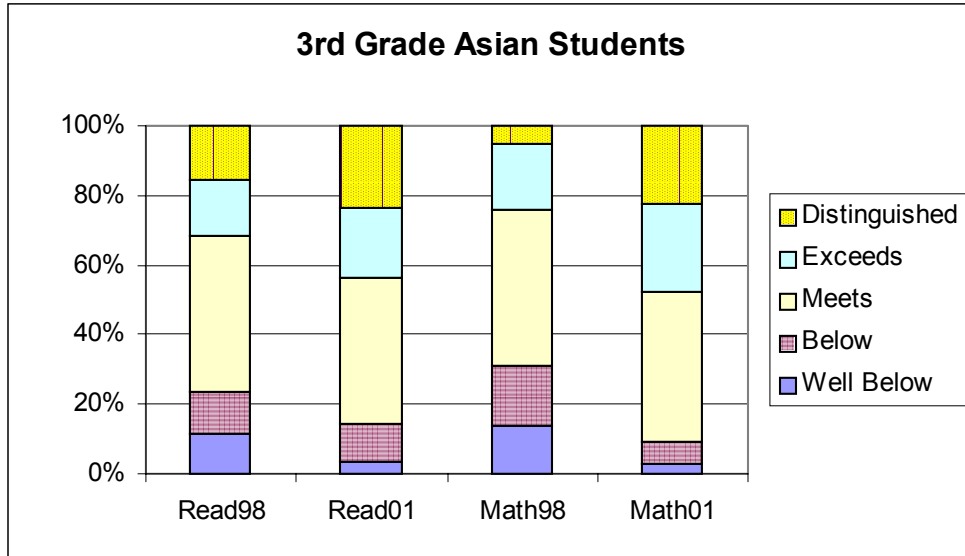
Graph 33



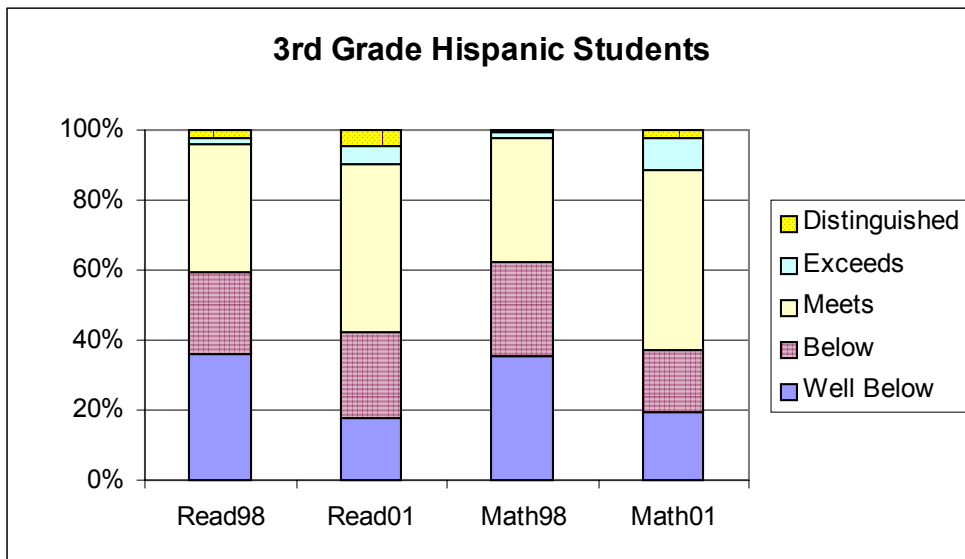
Graph 34



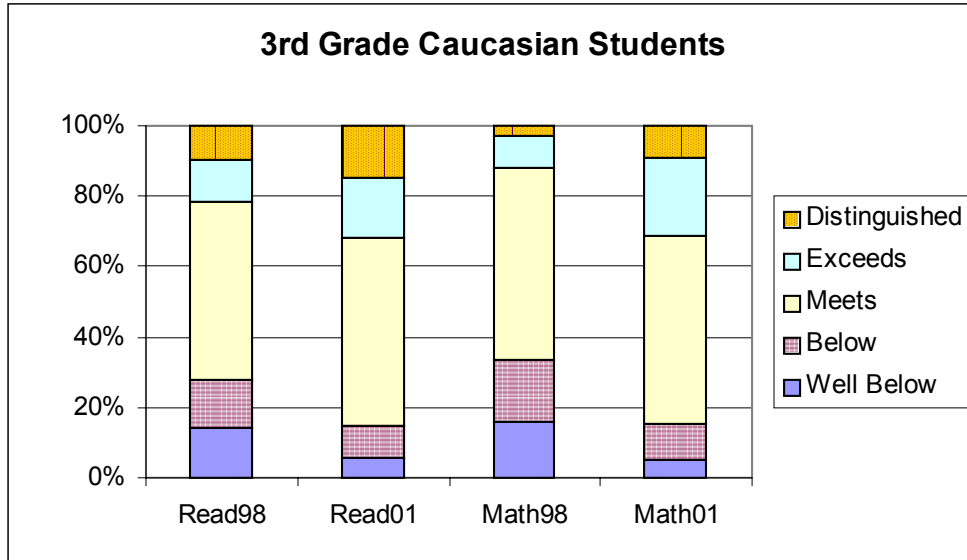
Graph 35



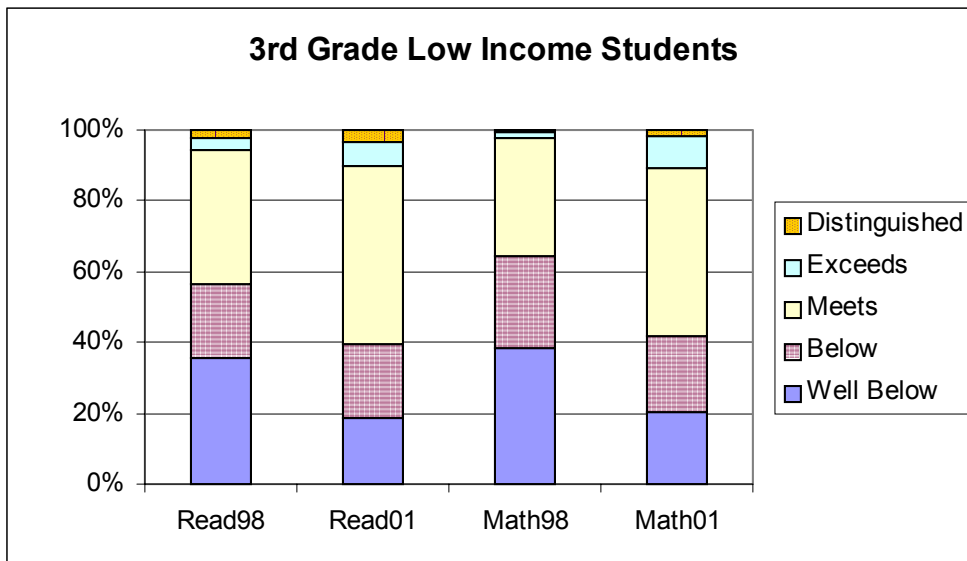
Graph 36



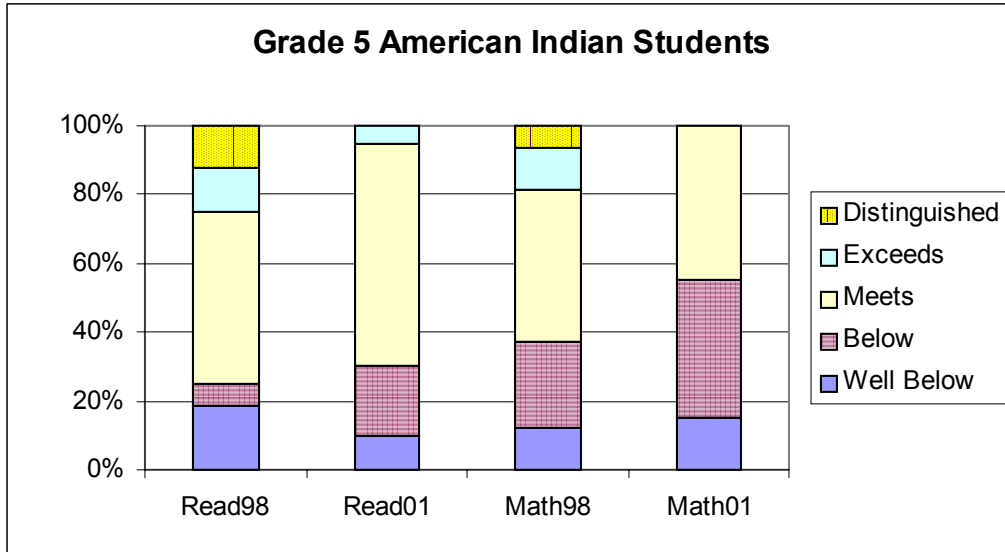
Graph 37



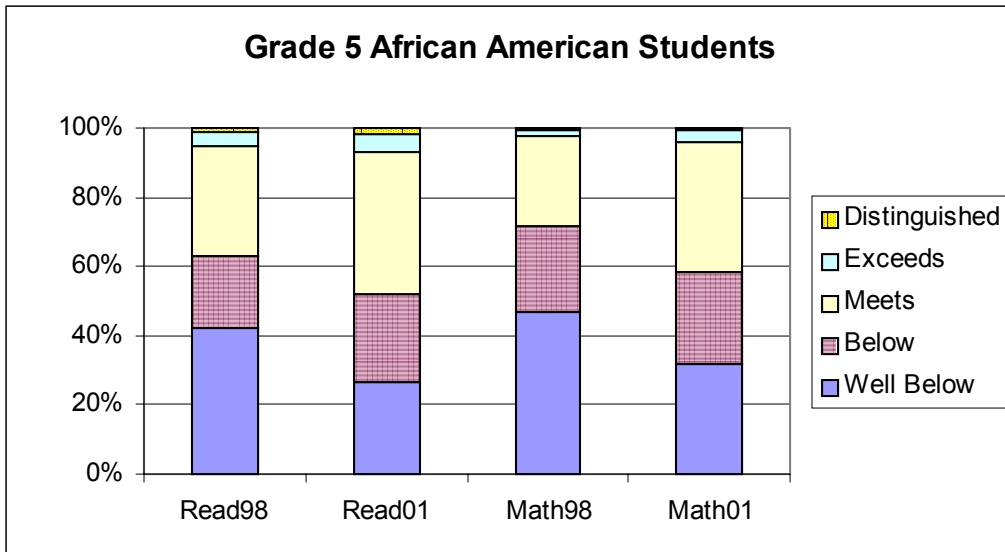
Graph 38



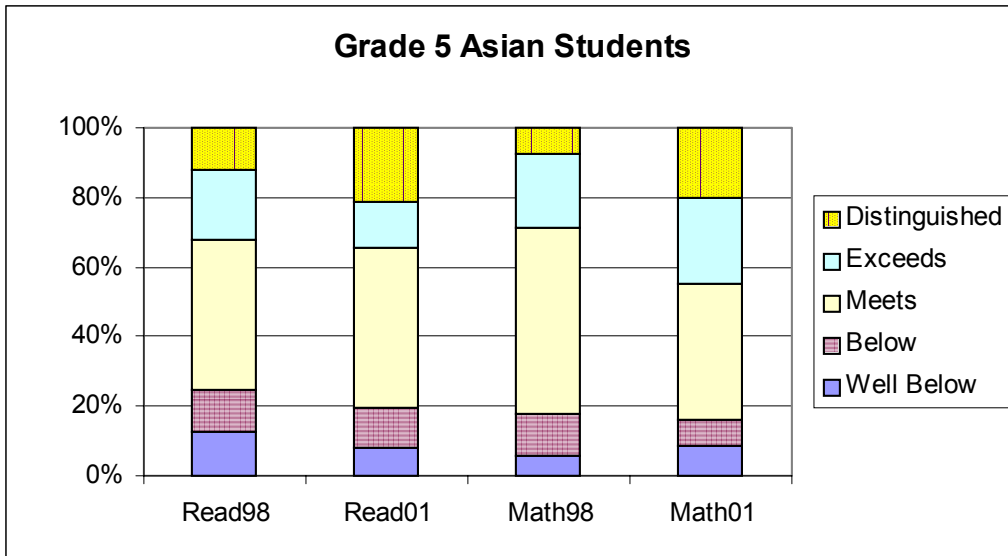
Graph 39



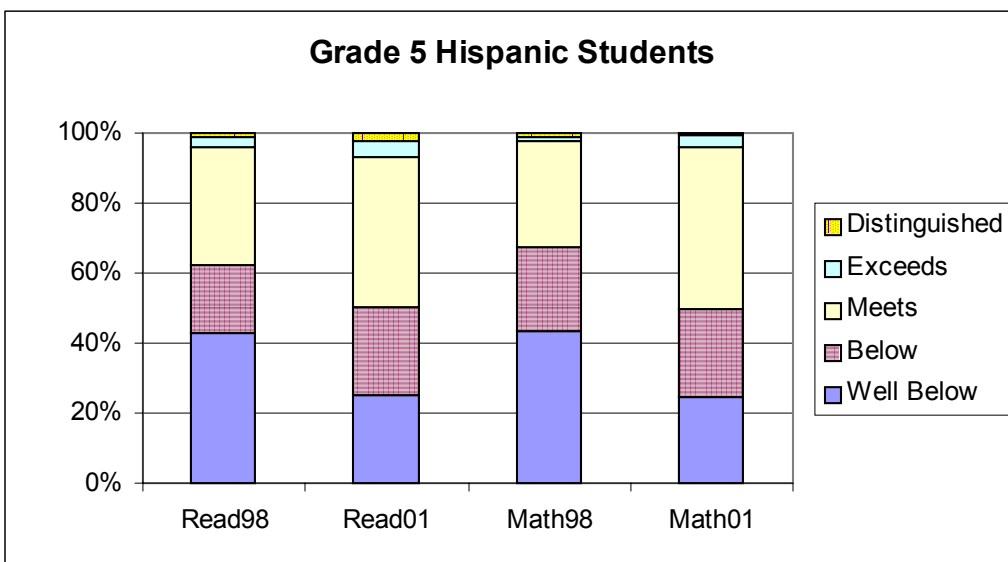
Graph 40



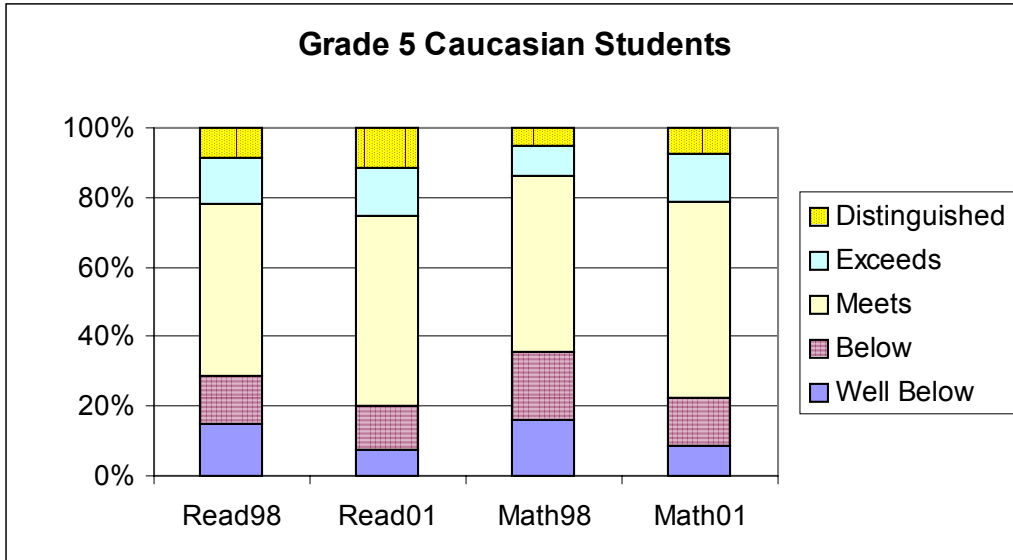
Graph 41



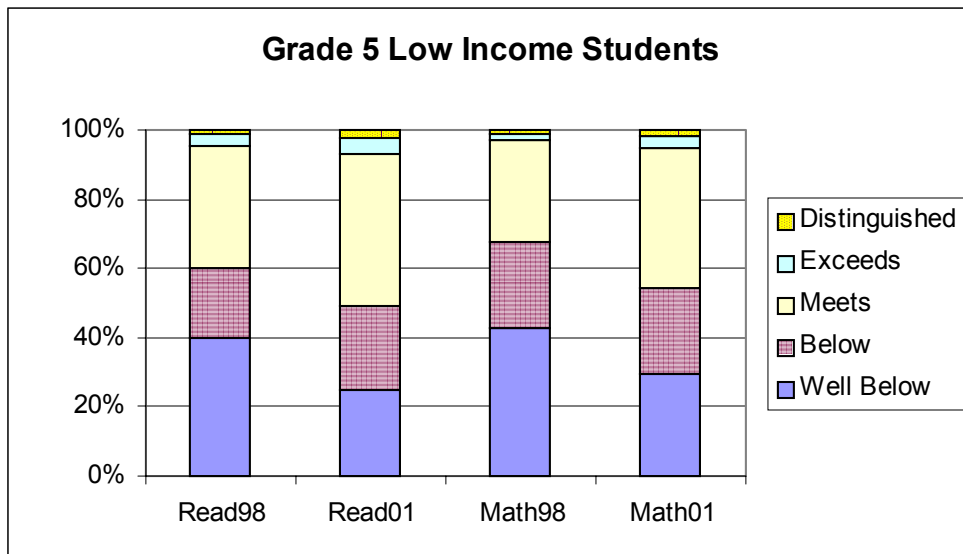
Graph 42



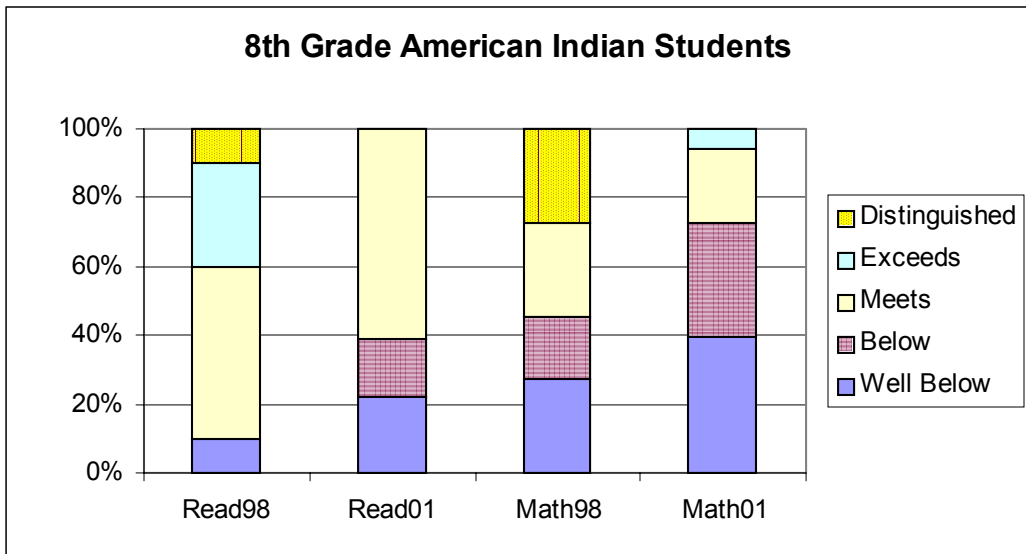
Graph 43



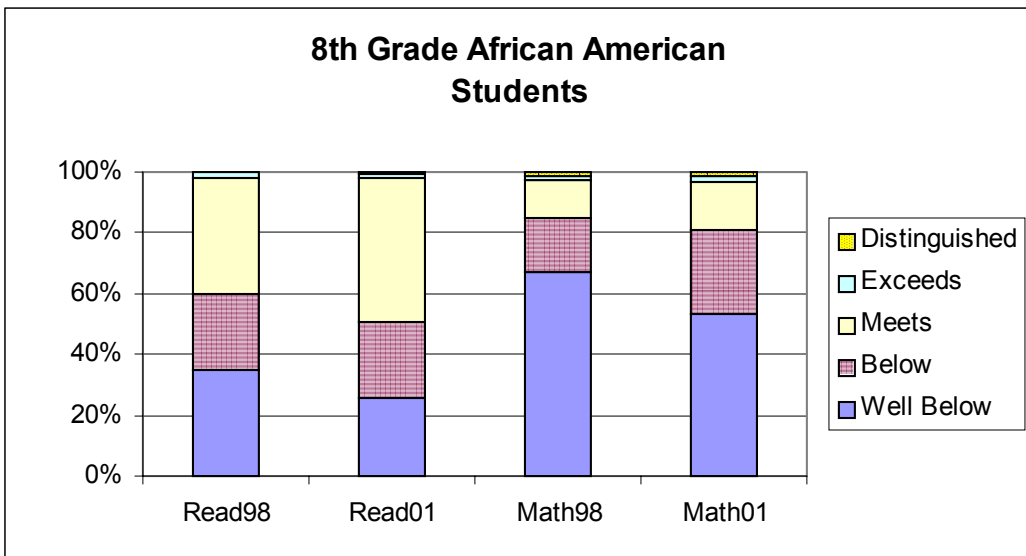
Graph 44



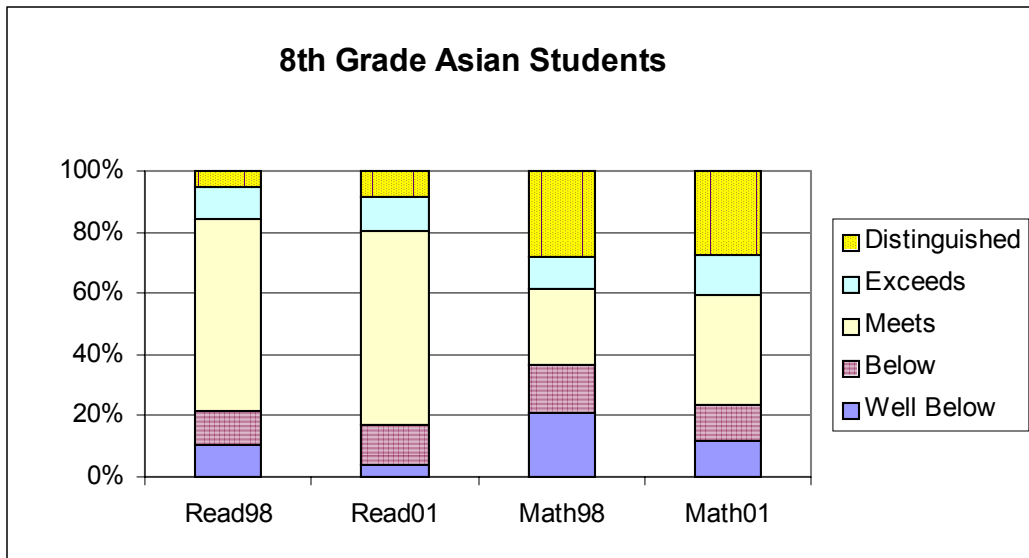
Graph 45



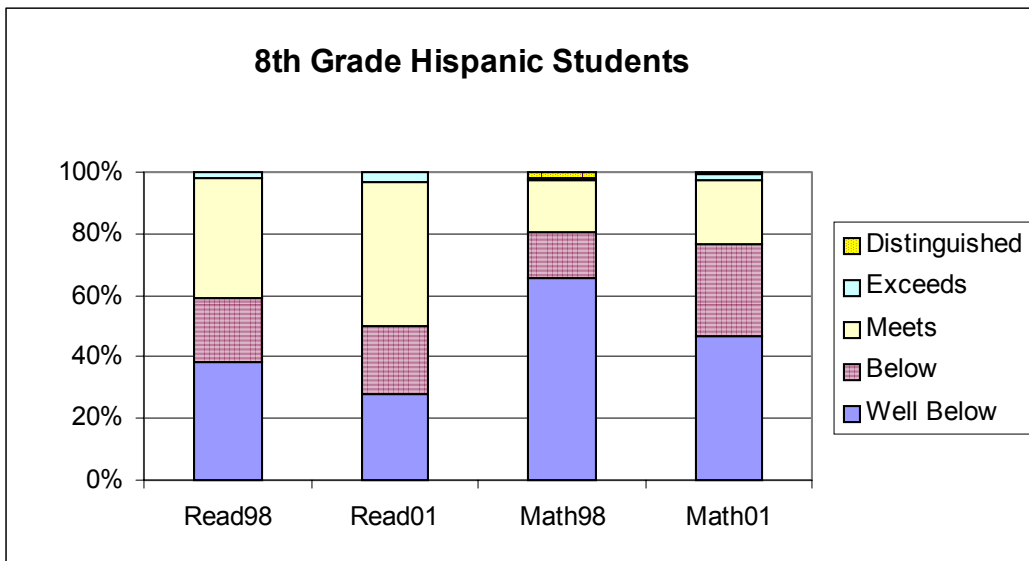
Graph 46



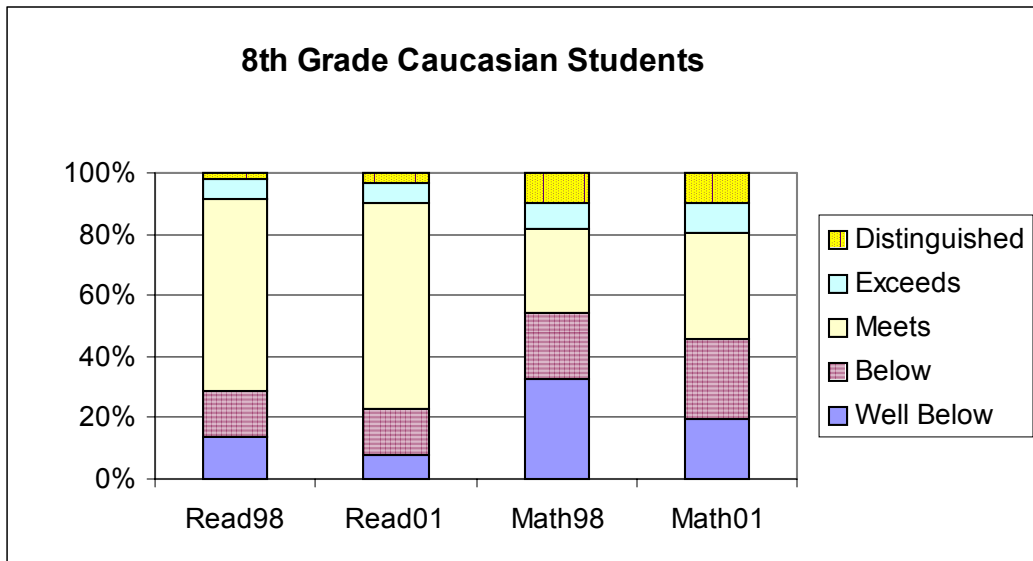
Graph 47



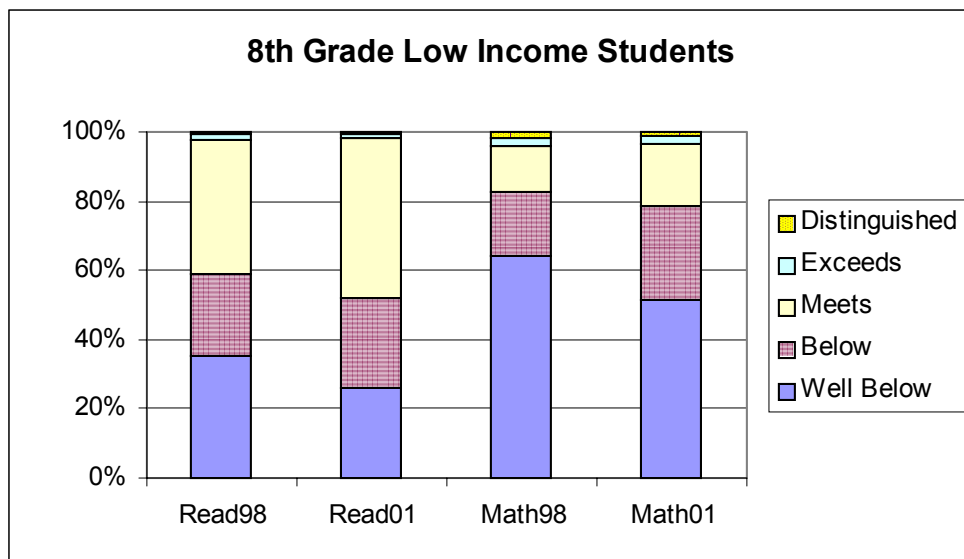
Graph 48



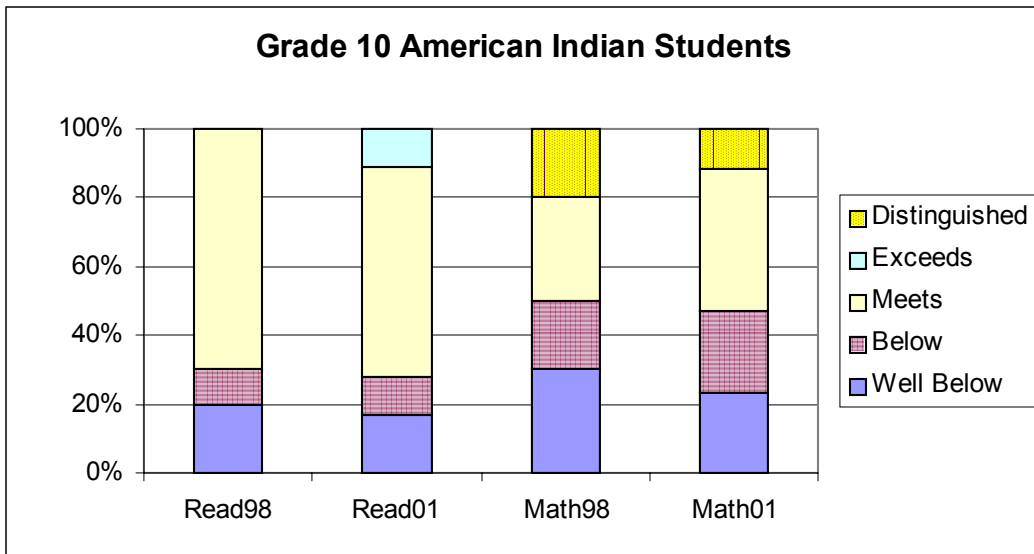
Graph 49



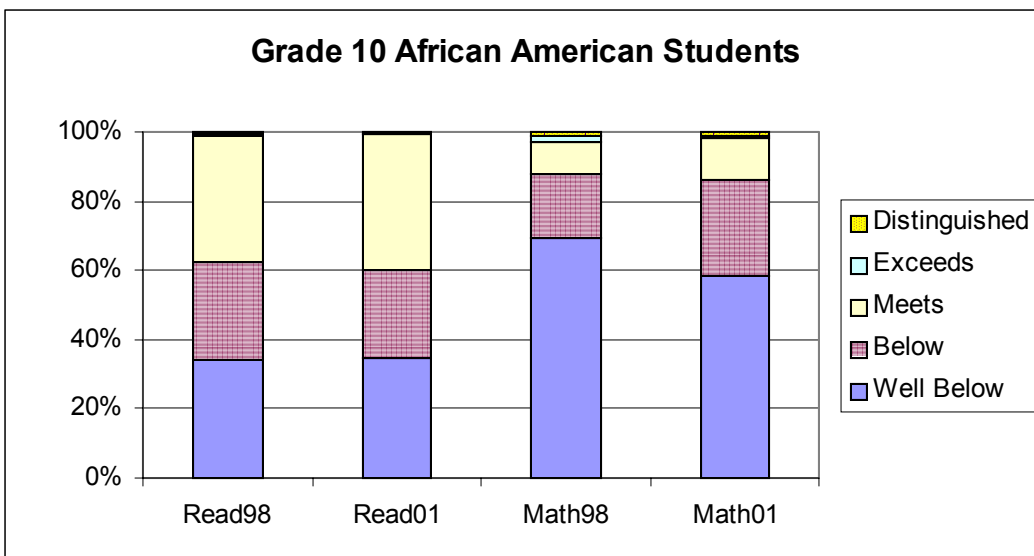
Graph 50



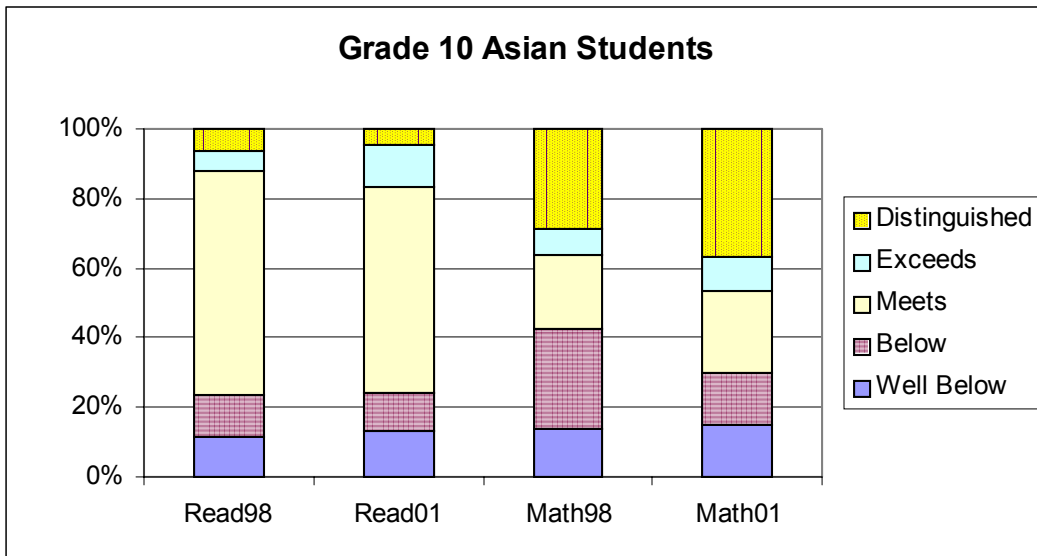
Graph 51



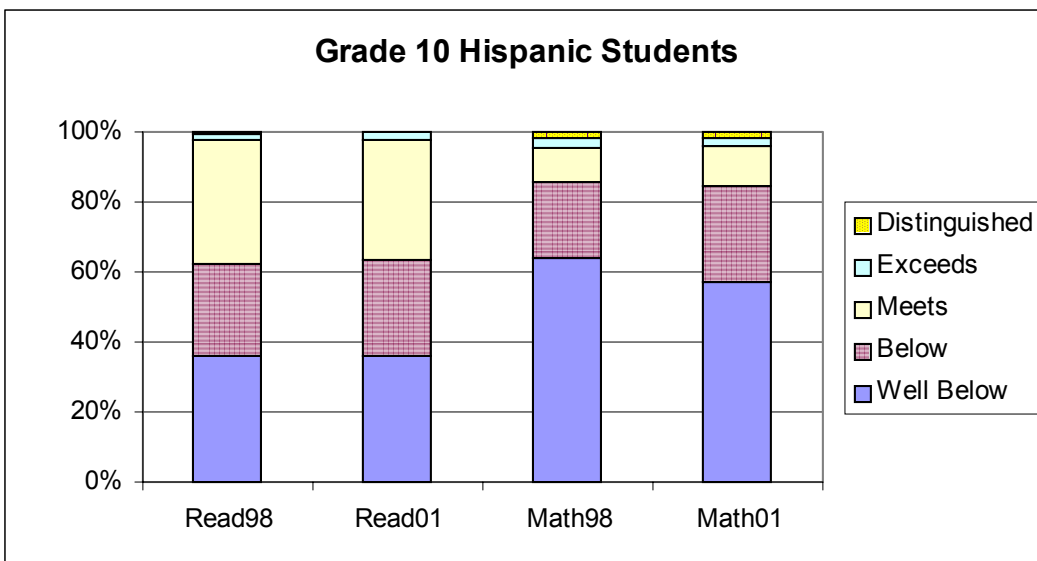
Graph 52



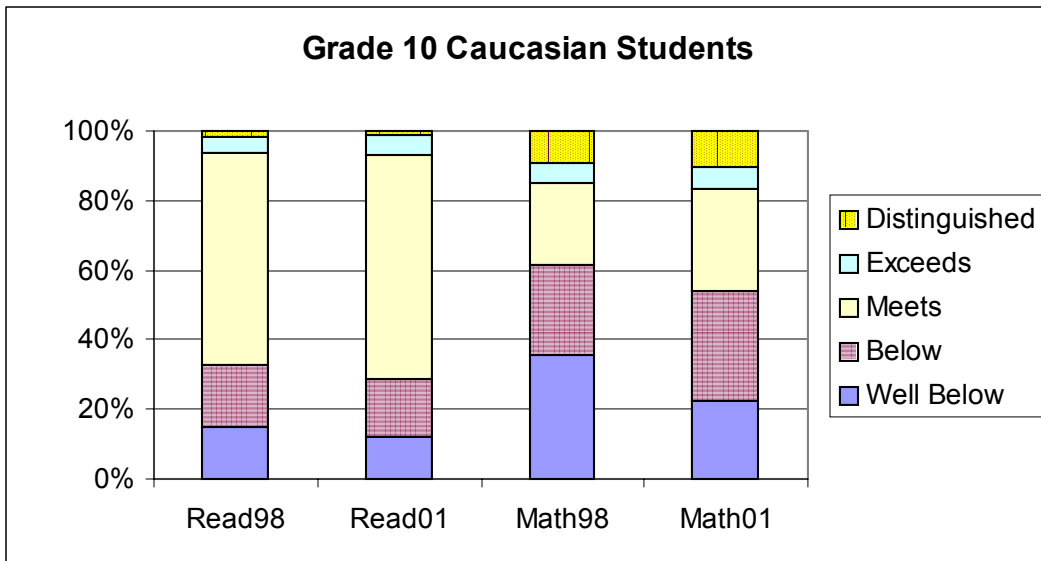
Graph 53



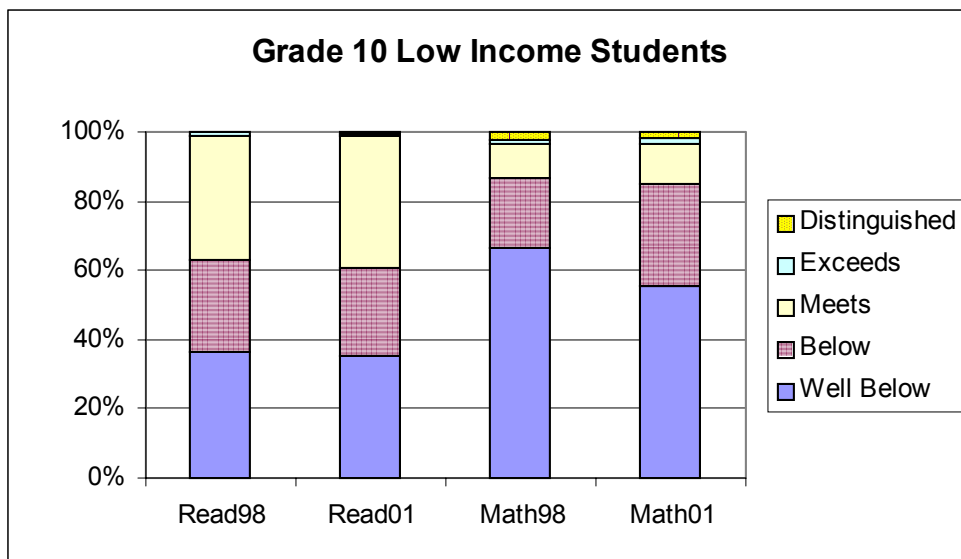
Graph 54



Graph 55



Graph 56



An examination of the percentage of students at each performance level in 1998 and 2001 indicated that for nearly all of the student groups the percentage of students at or above the standard increased over time.* Combined with the information provided above concerning the achievement gap among student groups, the findings indicated that the decreasing gaps in performance were the type to be desired.

The results further indicated that the changes in the percentage of students at or above the standard appeared to be larger in the elementary grades while the changes in percentages at the 10th grade level were very small. While there were increases in the percentage of students at or above the standard between 1998 and 2001, the total percentages of students at or above the standard in the most recent testing year ranged widely according to the student group, the content area, and the grade under consideration. Figures 5 and 6 display the percentage of each student group scoring at or above the standard in the area of reading and mathematics in 2001.

Figure 5

% of Students At or Above the Standard in Reading—Spring 2001				
	Grade 3	Grade 5	Grade 8	Grade 10
American Indian	76.4	70	61.1	72.2
African American	58.5	48	49.3	40.1
Asian	85.8	80.3	82.8	75.7
Hispanic	57.9	49.7	49.8	36.8
Caucasian	85.5	80.1	77.5	71.1
Low Income	60.3	50.6	48.2	39.4

* The results for American Indians were an exception to this finding but again the size of this group is generally very small and often not reported with the aggregated results of the DSTP.

Figure 6

% of Students At or Above the Standard in Mathematics—Spring 2001				
	Grade 3	Grade 5	Grade 8	Grade 10
American Indian	64.7	45	27.3	53
African American	53.6	41.8	19.4	13.7
Asian	90.9	84.1	76.4	69.9
Hispanic	62.8	50.5	23.5	15.2
Caucasian	84.8	77.5	54.1	46.2
Low Income	58.1	45.9	21.5	15.3

An examination of Figures 5 and 6 revealed that higher percentages of students were at or above the standard in the elementary grades when compared to the secondary grades. With the exception of 3rd and 5th grade Asian and Hispanic students, performance in the area of reading was superior to that of mathematics performance. The percentage of students at or above the mathematics standard was particularly low in the 10th grade.

SECTION 3: BEHAVIORAL EFFECTS

Research in the area of school discipline indicates that student misbehavior in the classroom is often a precursor for later school dropout and a variety of other negative social outcomes (i.e. poorer psychological adjustment and poorer academic self-concepts). Many strategies have been designed to address this problem with varying degrees of success.⁶

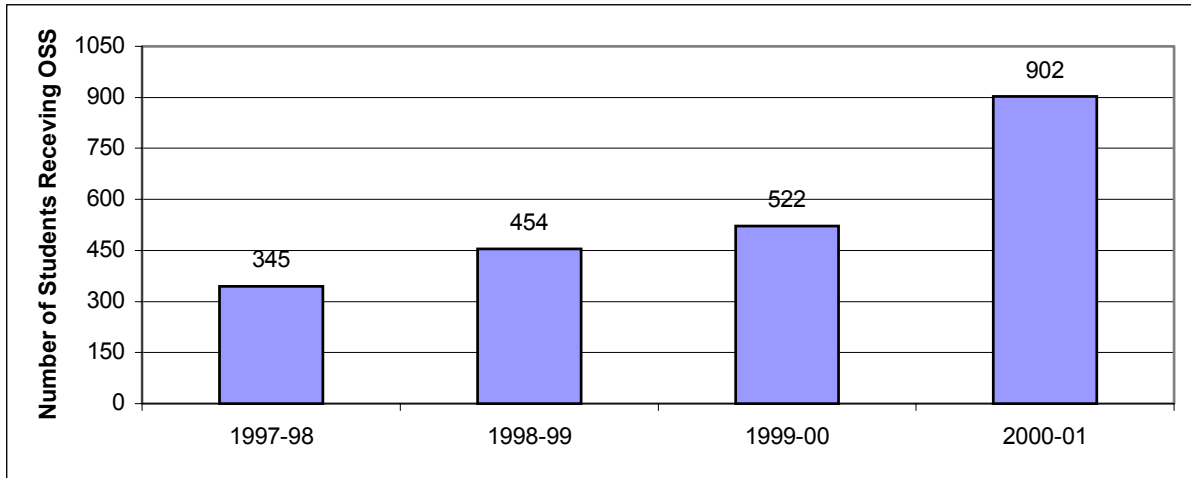
One of the least effective strategies for addressing student misbehavior involves the use of out-of-school suspension. Removing a misbehaving student from the school prevents the student from learning and may inadvertently increase misbehavior as a result of the frustration a student may experience from falling further behind in his or her school work. In addition, there is some research to indicate that out-of-school suspension is not an effective deterrent for future misbehavior in that large numbers of suspended students are often repeat offenders.⁷ In addition to being an ineffective discipline strategy, research also indicates that the strategy is often disproportionately based on race, socio-economic status, gender, and disability.⁸

The link of achievement to student behavior was investigated by examining the incidence of out-of-school suspension and its relationship to performance on the DSTP. In addition, because previous research has demonstrated a link between student demographics and incidences of suspensions, the demographic characteristics of students given out-of-school suspension were also examined.

ELEMENTARY COHORT SUSPENSIONS: STUDENT DEMOGRAPHICS

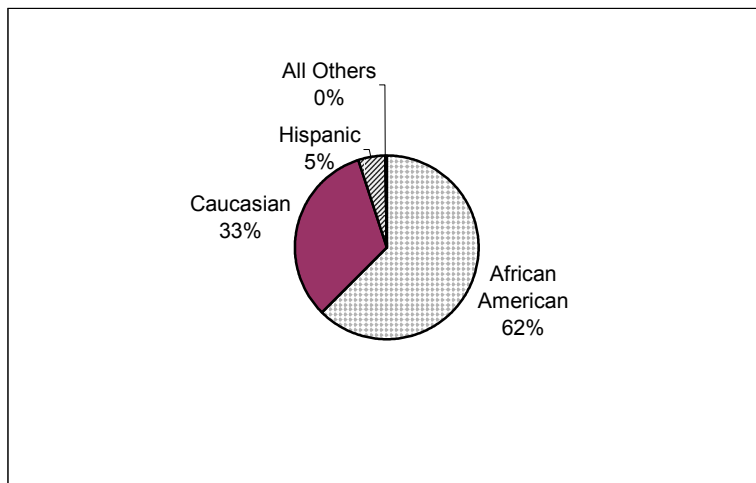
The total number of students suspended in the elementary cohort for each of the four years of study appears in Graph 57. The number of students suspended increased during each year of the study with the highest number reported during the 2000-01 school year. This increased rate over time is disconcerting considering that some attrition has occurred in the elementary cohort over time, in which case the number of students in the cohort has decreased while the number of students suspended has increased.

Graph 57: Annual Number of Out-of-School Suspensions for the Elementary Cohort



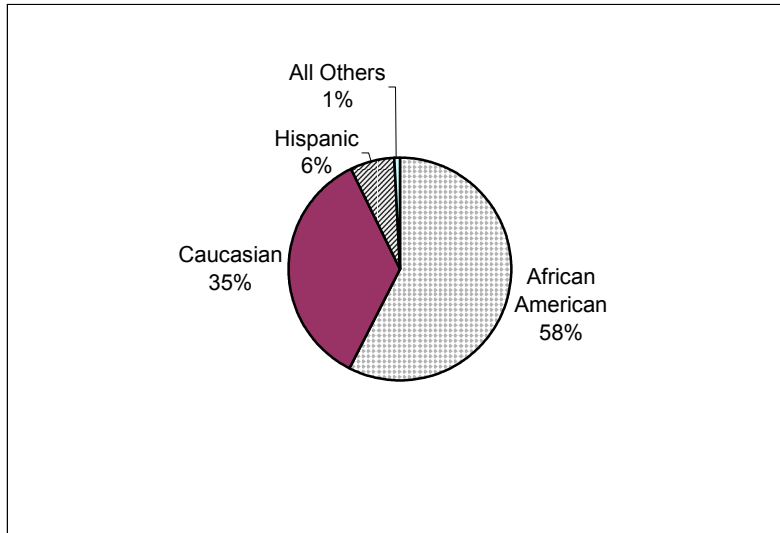
Of the 522 students receiving out-of-school suspensions (OSS) in 1999-00, 77.2% were male and 70.5% were classified as low income. The ethnicity of the students receiving OSS during the 1999-00 school year appears in Graph 58.

Graph 58: Elementary Cohort: Ethnicity of Students Receiving OSS in 1999-00



An examination of the demographic characteristics of students receiving out-of-school suspension (OSS) in 2000-01 revealed that 76.7% were male and 66.4% were low income students. The ethnicity of the students receiving OSS during the 2000-01 school year appears in Graph 59.

Graph 59: Elementary Cohort: Ethnicity of Students Receiving OSS in 2000-01

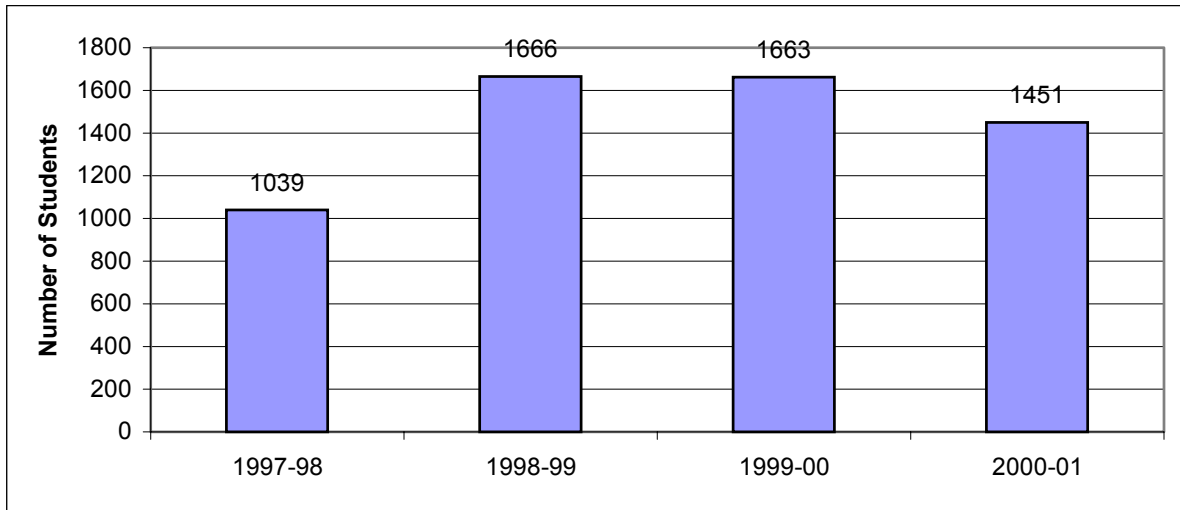


Logistic regression analyses were again conducted to examine the relationship between student demographics and the odds of being suspended. In keeping with the findings of last year's longitudinal report and previous research on out-of-school suspension, the results of these analyses revealed that income status, gender, and minority status were statistically significant predictors of suspensions given during the 1999-00 and 2000-01 school years. The odds of being suspended were more than 3 times greater for males compared to females, more than 2 times larger for low income students than non-low income students, and more than 2 times larger for African American students than Caucasians.

SECONDARY COHORT SUSPENSIONS: STUDENT DEMOGRAPHICS

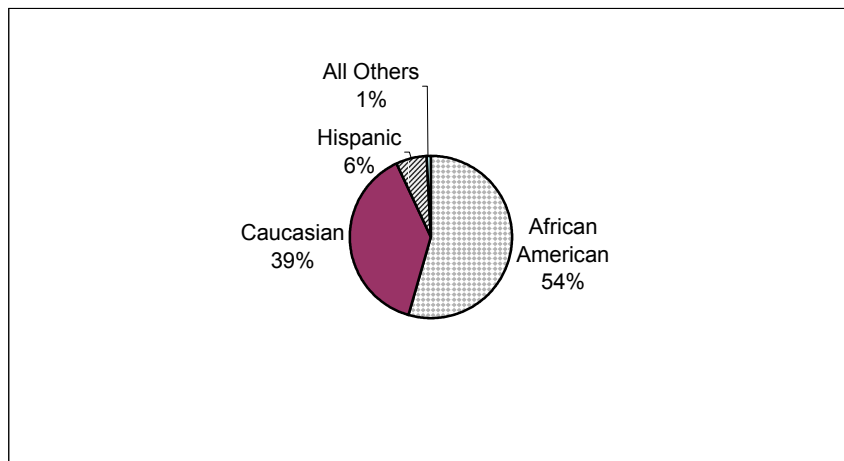
The total number of students suspended in the secondary cohort for each of the four years of study appears in Graph 60. Unlike the elementary cohort, the number of suspensions in the secondary cohort has not increased over time but leveled off between 1998 and 1999 with fewer suspensions during the 2000-01 school year. As with the elementary cohort some attrition has occurred within the cohort.

Graph 60: Annual Number of Out-of-School Suspensions for the Secondary Cohort



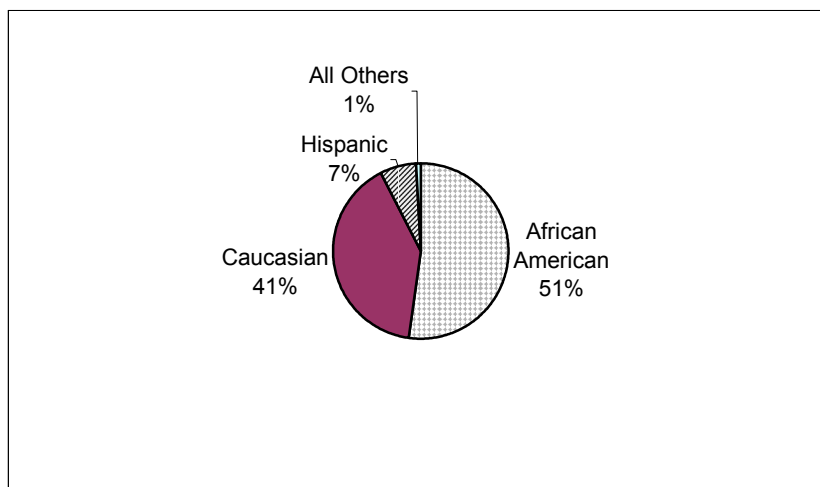
An examination of the demographic characteristics of students receiving out-of-school suspension (OSS) in 1999-00 revealed that 64.2% were male and 56.9% were low income students. The ethnicity of the students receiving OSS during the 1999-00 school year appears in Graph 61.

Graph 61: Secondary Cohort: Ethnicity of Students Receiving OSS in 1999-00



Of the 1451 students suspended during the 2000-01 school year, 60.1% were male and 46.5% were low income students. The ethnicity of the students receiving OSS during the 2000-01 school year appears in Graph 62.

Graph 62: Secondary Cohort: Ethnicity of Students Receiving OSS in 2000-01



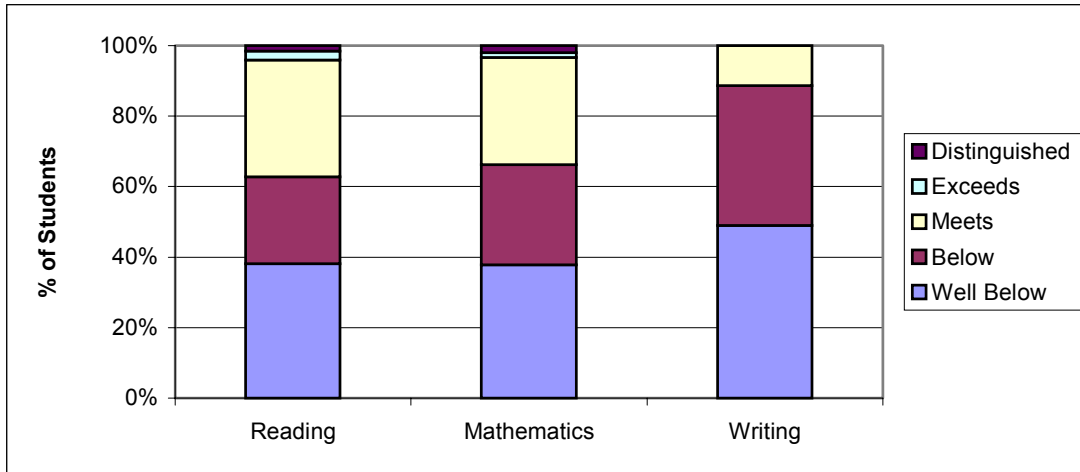
Logistic regression analyses were again conducted to examine the relationship between student demographics and the odds of being suspended. In keeping with the findings of last year’s longitudinal report and previous research on out-of-school suspension, the results of this analysis revealed that income status, gender, and minority status were statistically significant predictors of suspensions given during the 1999-00 and 2000-01 school years. The odds of being suspended were more than 1.6 times greater for males compared to females, more than 2 times larger for low income students than non-low income students, and more than 2 times larger for African Americans than Caucasians.

SUSPENSION AND DSTP PERFORMANCE

In addition to examining the demographic characteristics of suspended students the current investigation also examined the relationship between suspension and student performance on the DSTP. Because most of the elementary and secondary cohorts completed the 5th and 8th grade version of the DSTP in the Spring of 2000, the performance of students suspended during the 1999-00 school year was examined.

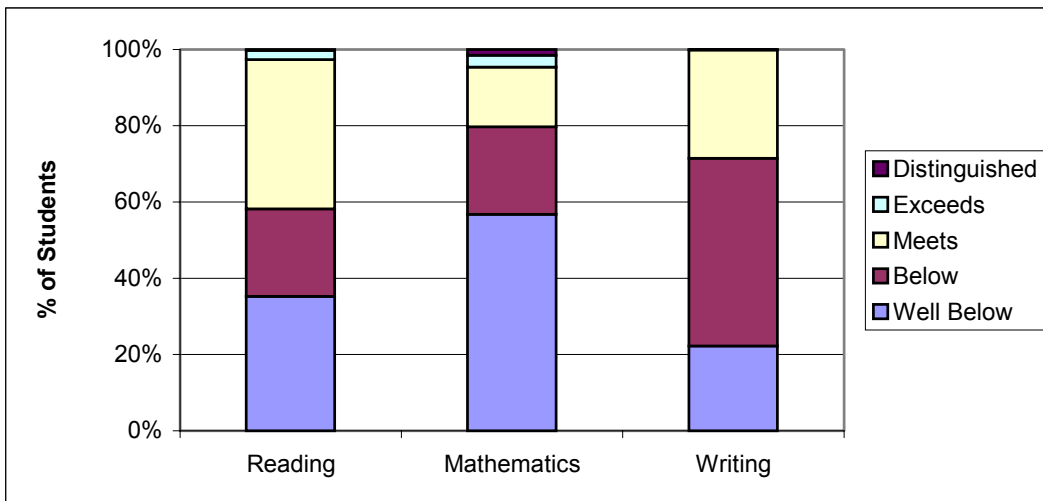
Graph 63 presents the results of this examination within the elementary cohort. According to the data, fewer than 40% of the elementary students suspended during the 1999-00 school year met or exceeded the state standards in reading and mathematics. Fewer than 15% of these students performed at or above the standard in writing.

Graph 63: Elementary Cohort: Spring 2000 DSTP Performance for Students Receiving OSS in 1999-00



Among the secondary cohort, slightly more than 40% of the students suspended during the 1999-00 school year were at or above the standard in reading. The results also revealed that fewer than 20% of the suspended students were at or above the standard in mathematics and less than 30% were at or above the standard in writing (See Graph 64).

Graph 64: Secondary Cohort: Spring 2000 DSTP Performance for Students Receiving OSS in 1999-00



The purpose of examining the relationship between DSTP performance and out-of-school suspensions was not to argue a cause and effect relationship between the two but to

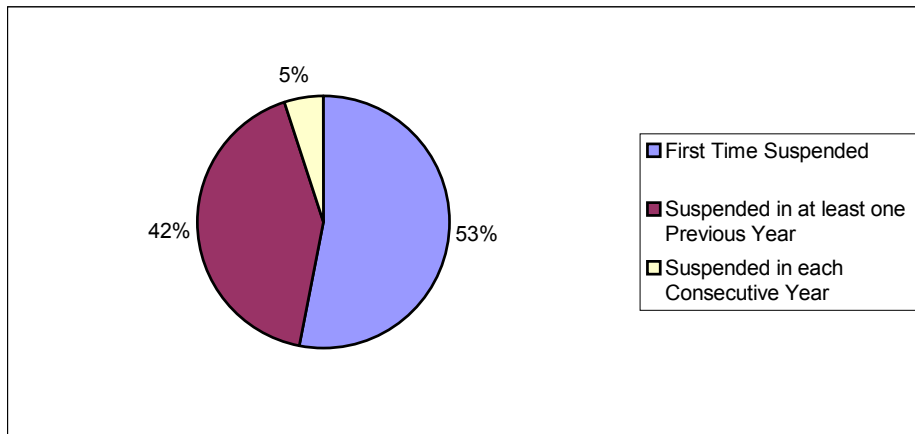
highlight the fact that suspension and poor academic performance may go hand-in-hand. What is clear is that students with higher odds of being suspended are also those students who have historically been under-performers academically. It is unlikely that removing students from school will help matters improve.

SUSPENSION HISTORIES FOR THE COHORTS

Research clearly indicates that out-of-school suspension is not an effective strategy for student remediation. As a further demonstration, a cursory examination of the extent to which suspension may act as a deterrent to future misbehavior was examined using the elementary and secondary cohorts. The “suspension history” of students suspended during the 2000-01 school year was investigated by observing whether a student had been suspended in a previous year of study.

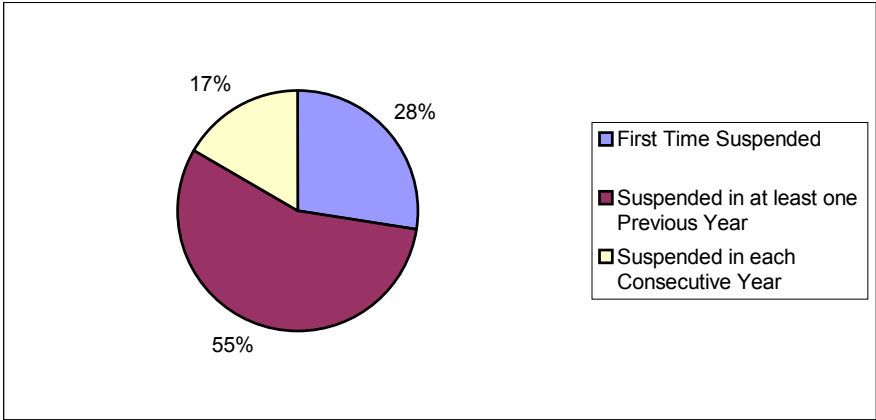
An examination of the data revealed that slightly less than half of the elementary students suspended during the 2000-01 school year had been previously suspended in another year with 5% of students being suspended once during each year of study (See Graph 65).

Graph 65: Suspension History for Students Suspended during the 2000-01 School Year (Elementary Cohort)



Among the secondary cohort as many as 55% of the students suspended in 2000-01 had been previously suspended and as many as 17% had been suspended once during each year of study (See Graph 66). These results suggest that out-of-school suspension may not be an effective deterrent to future misbehavior.

Graph 66: Suspension History for Students Suspended during the 2000-01 School Year (Secondary Cohort)



SECTION 4: RETENTION

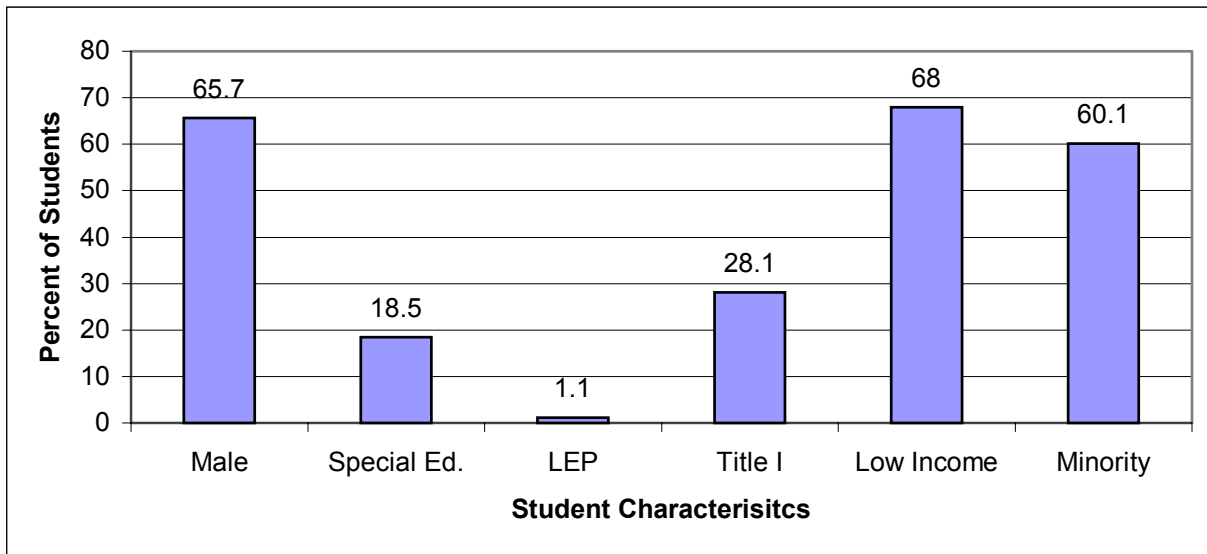
Research on high stakes assessment systems reveals that retention rates often increase when such systems are implemented. There is also a great deal of evidence to suggest that grade retention as typically practiced is often harmful and more likely to occur among certain student groups. One study examining the impact of retention on dropout rates revealed that when a student repeated a grade, his/her chances of later dropping-out increased by 20 to 40 percent.⁹

For these reasons, the incidence of retention and the demographic characteristics of retained students were examined. The relationship of retention to DSTP performance was also examined with the available data.

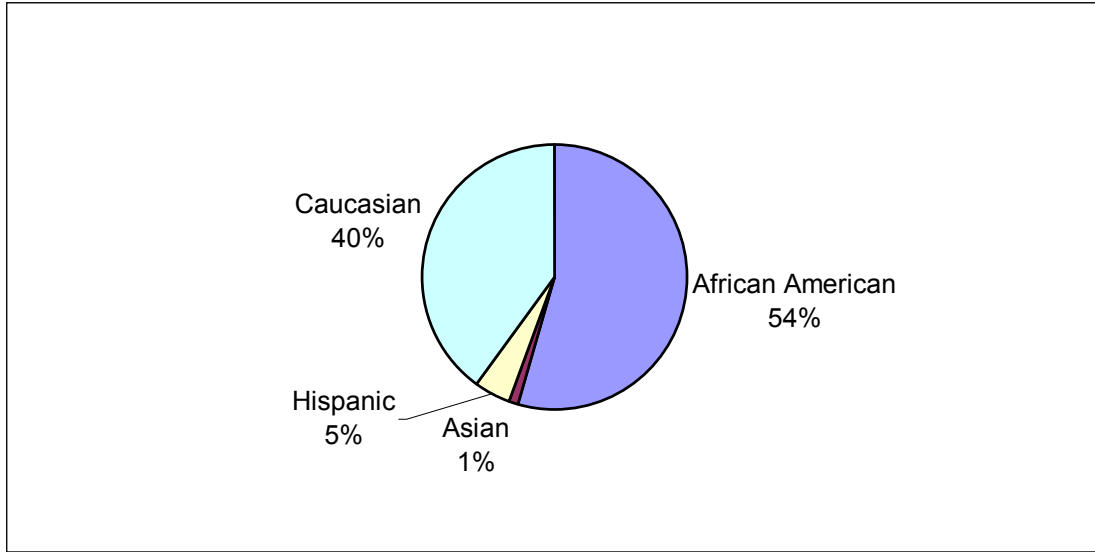
ELEMENTARY COHORT: DEMOGRAPHIC CHARACTERISTICS OF RETAINEES

The demographic characteristics of students in the elementary cohort who were enrolled in grade 4 during the 1998-99 **and** 1999-00 school years (N=178) were as follows:

Graph 67: Demographic Information for Students retained in Grade 4

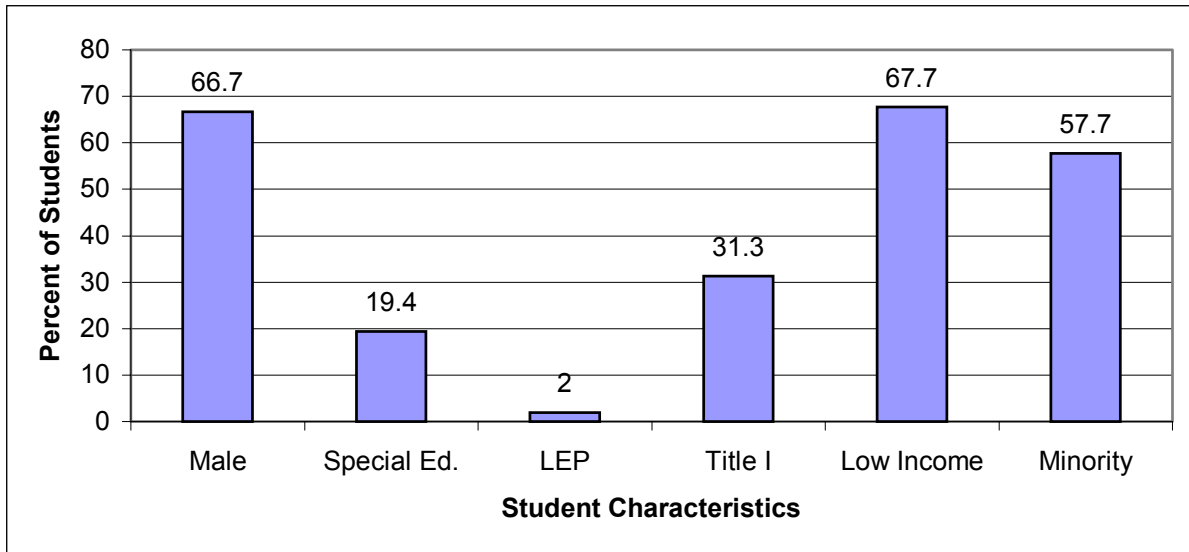


Graph 68: Ethnicity of Students Retained in Grade 4

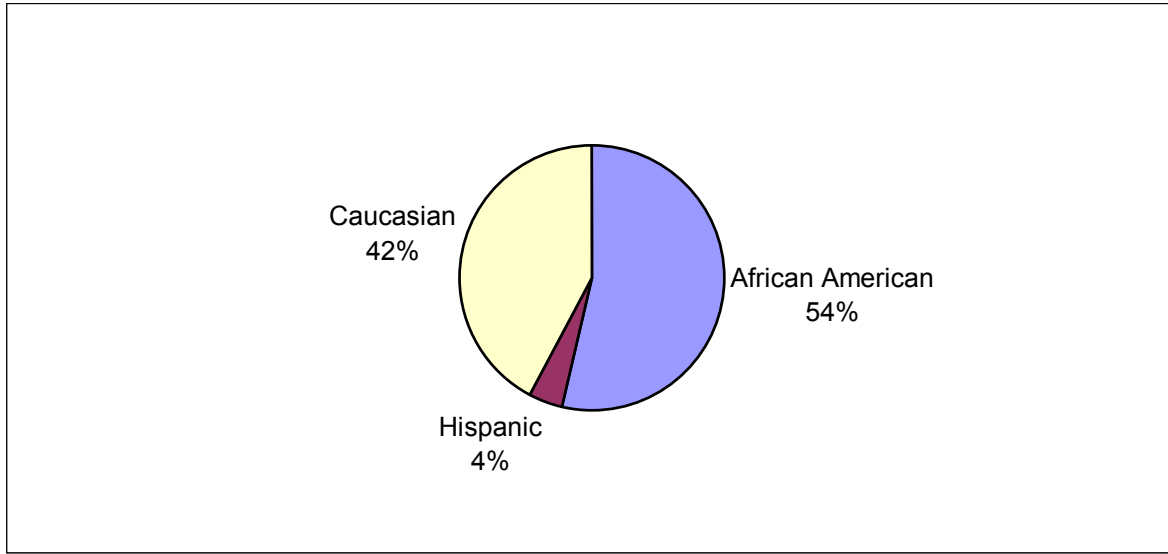


The demographic characteristics of students in the elementary cohort who were enrolled in grade 5 during the 1999-00 **and** 2000-01 school year (N=201) were as follows:

Graph 69: Demographics Information for Students Retained in Grade 5



Graph 70: Ethnicity of Students Retained in Grade 5

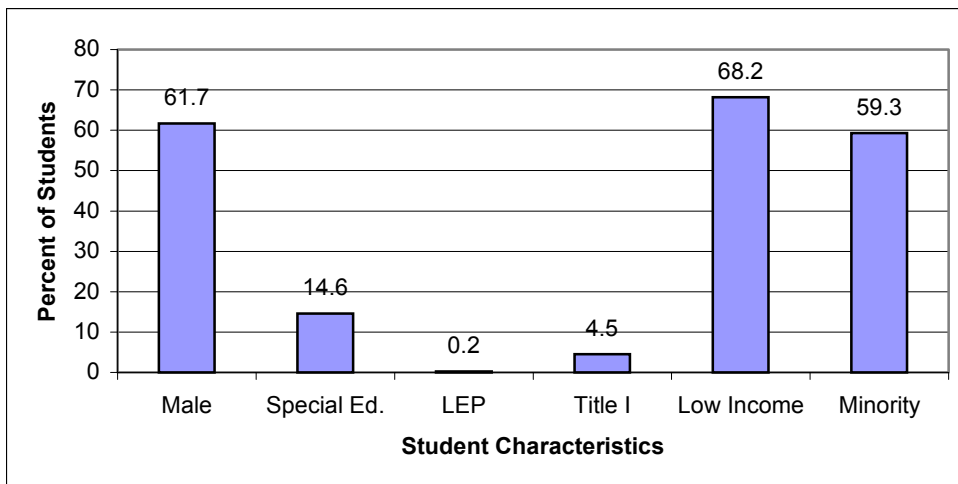


Results of logistic regression analyses using race, income status, and gender to predict retention indicated that each of these demographic characteristics were useful predictors. The odds of being retained were more than 1.7 times higher for males compared to females, more than 1.5 times greater for African Americans compared to Caucasians, and more than 2.3 times greater for low income students than non-low income students. These results replicate previous research and the results of last year’s study.

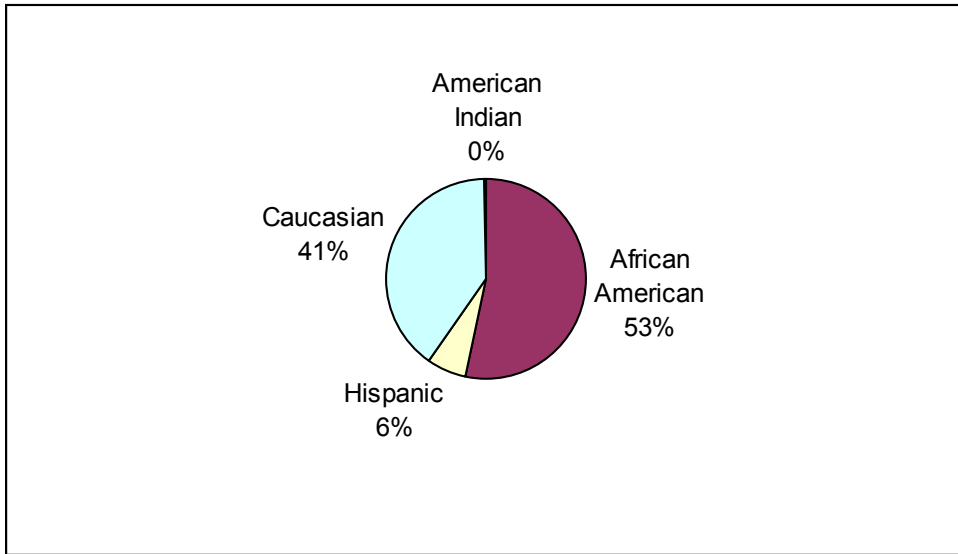
SECONDARY COHORT: DEMOGRAPHIC CHARACTERISTICS OF RETAINEES

The demographic characteristics of students in the secondary cohort who were enrolled in grade 7 during the 1998-99 **and** 1999-00 school year (N=494) were as follows:

Graph 71: Demographics of Students Retained in Grade 7

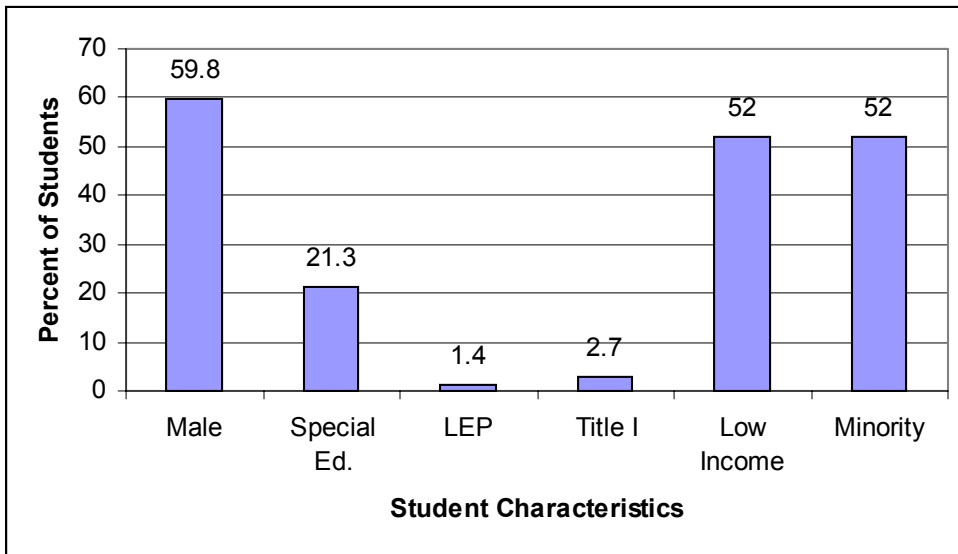


Graph 72: Ethnicity of Students Retained in Grade 7

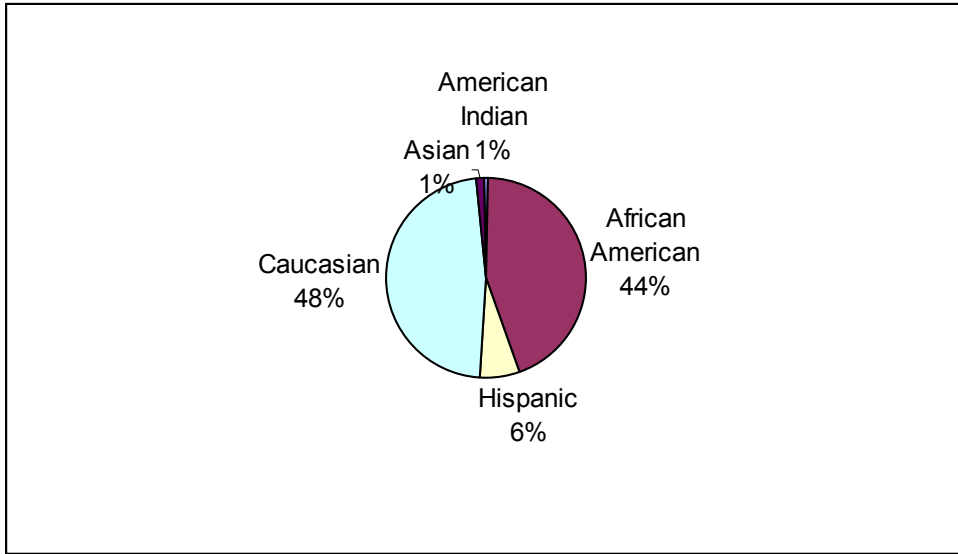


The demographic characteristics of students in the secondary cohort who were enrolled in grade 8 in 1999-00 **and** 2000-01 (N=296) were as follows:

Graph 73: Demographics of Students Retained in Grade 8



Graph 74: Ethnicity of Students Retained in Grade 8



The results of logistic regression analyses using race, income status, and gender to predict retention indicated that each of these demographic characteristics were useful predictors. The odds of being retained were more than 1.5 times higher for males, more than 1.5 times greater for minority students, and more than 2 times greater for low income students. These results replicate previous research and the results of last year's study.

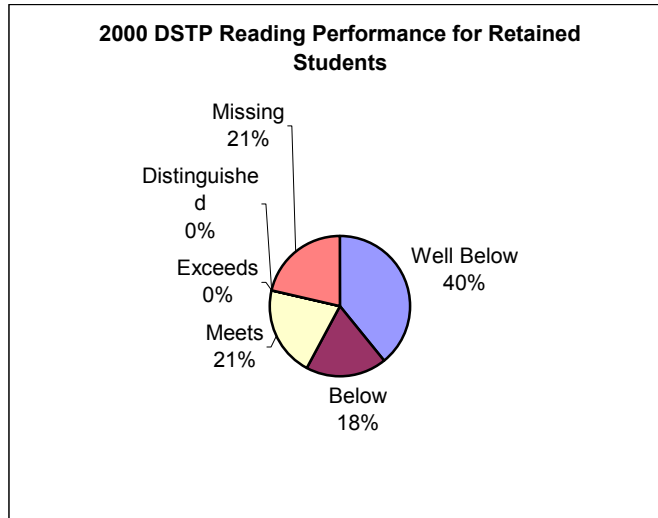
RETENTION AND DSTP PERFORMANCE

In addition to examining the demographic characteristics of retained students, the DSTP performance levels of these students were also examined. Students retained in 5th grade or 8th grade following the Spring 2000 administration of the DSTP were included in these analyses.

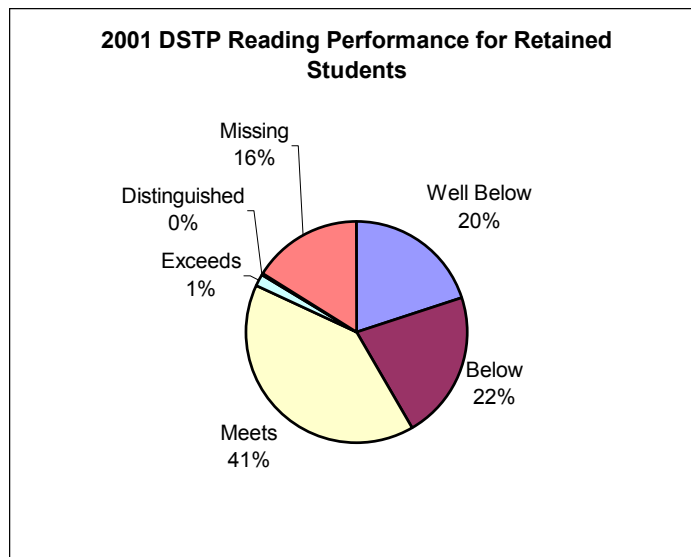
ELEMENTARY COHORT

Students who were enrolled in 5th grade in both the 1999-00 and 2000-01 school year could have completed the 5th grade version of the DSTP in the spring of 2000 and the spring of 2001. The DSTP performance levels of students who were eventually retained in grade 5 were as follows:

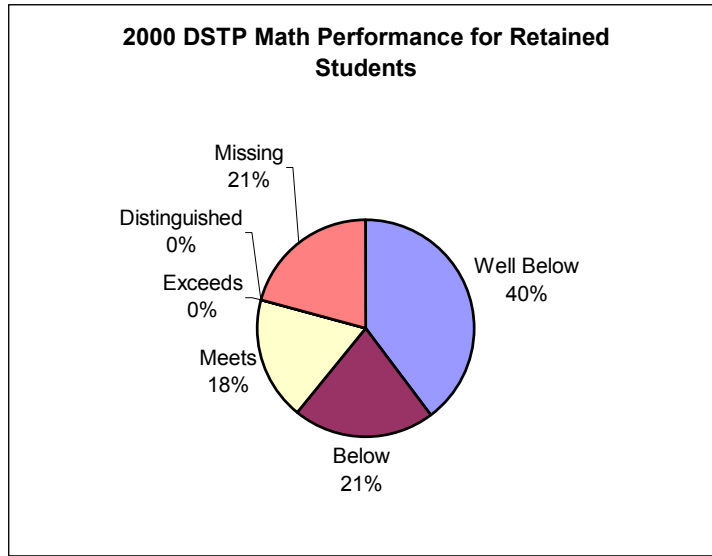
Graph 75



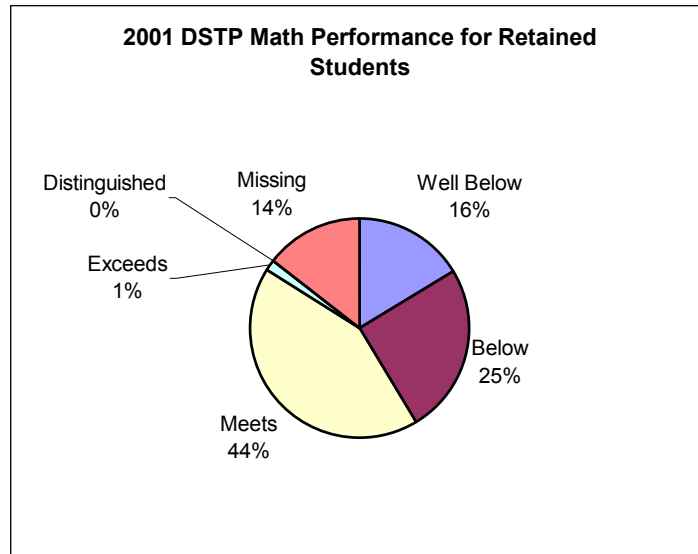
Graph 76



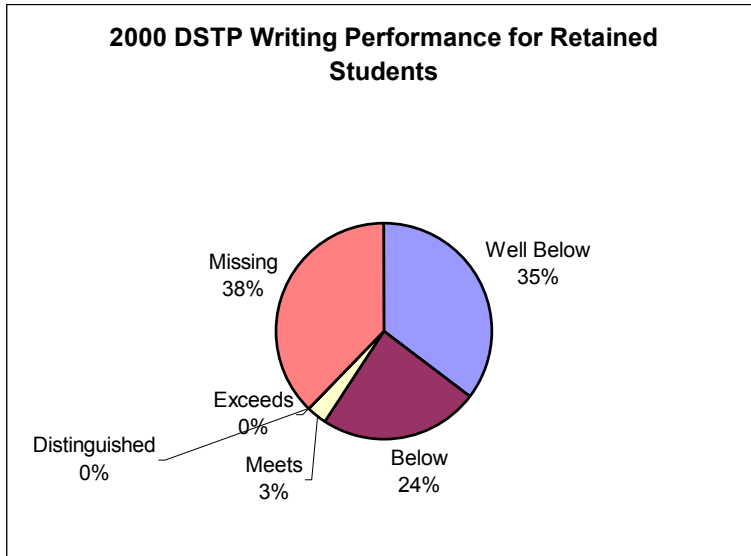
Graph 77



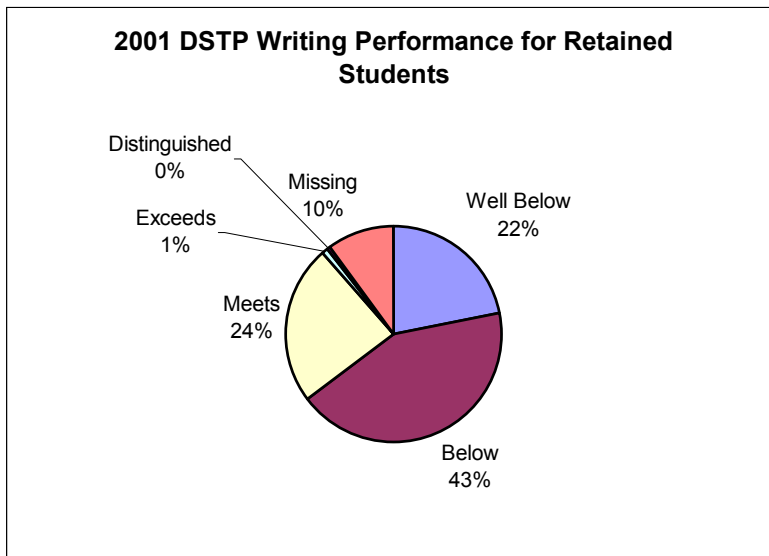
Graph 78



Graph 79



Graph 80

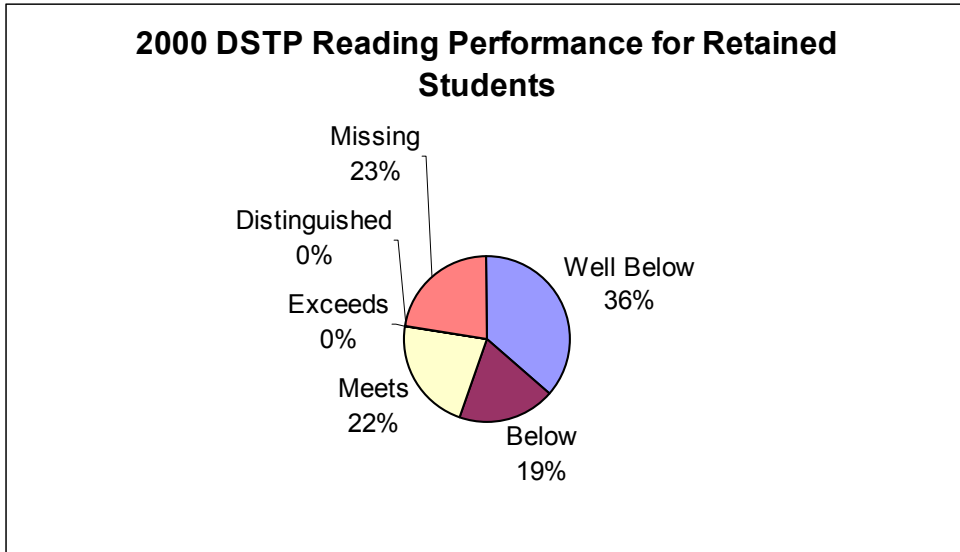


An examination of the data revealed that larger percentages of retained students were able to meet or exceed the reading and math standard when taking the DSTP during the Spring 2001 administration. However, less than one-third of the retained students were able to meet the writing standard. In the areas of reading and mathematics, more than one-third of students who were retained in grade 5 performed below the standard in 2001. In the area of writing, 65% of the retained students were below the standard in 2001.

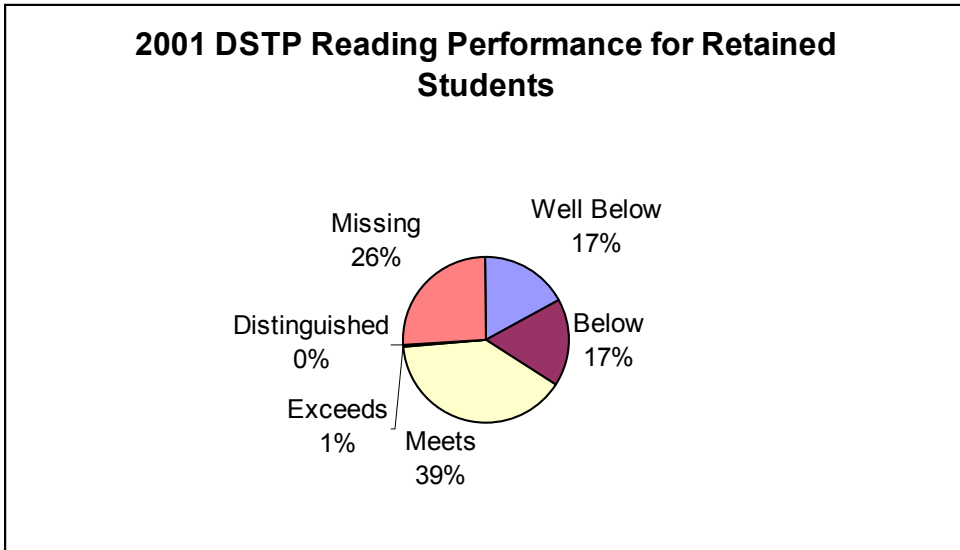
SECONDARY COHORT: RETENTION AND DSTP PERFORMANCE

Students who were enrolled in 8th^h grade in both the 1999-00 and 2000-01 school year could have completed the 8th^h grade version of the DSTP in the spring of 2000 and the spring of 2001. The DSTP performance levels of students who were eventually retained in grade 8 were as follows:

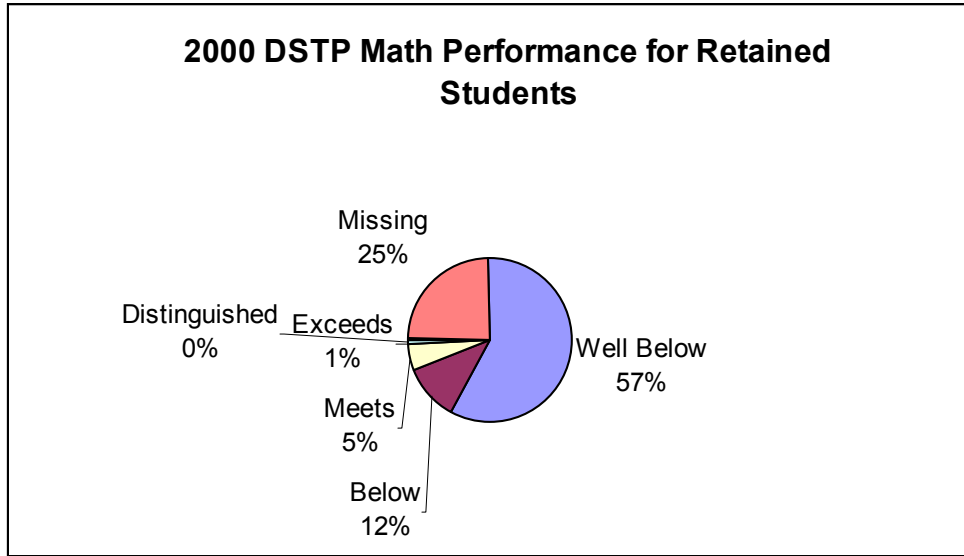
Graph 81



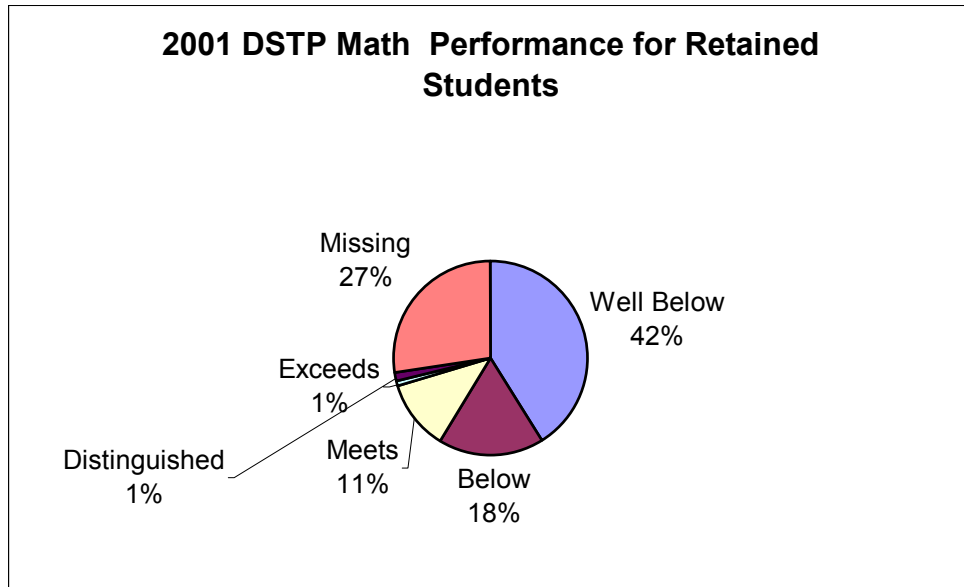
Graph 82



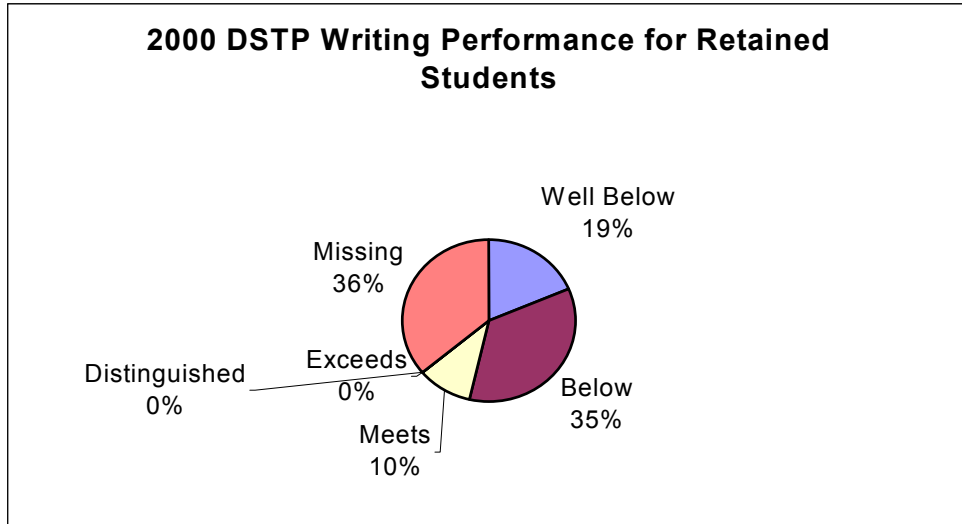
Graph 83



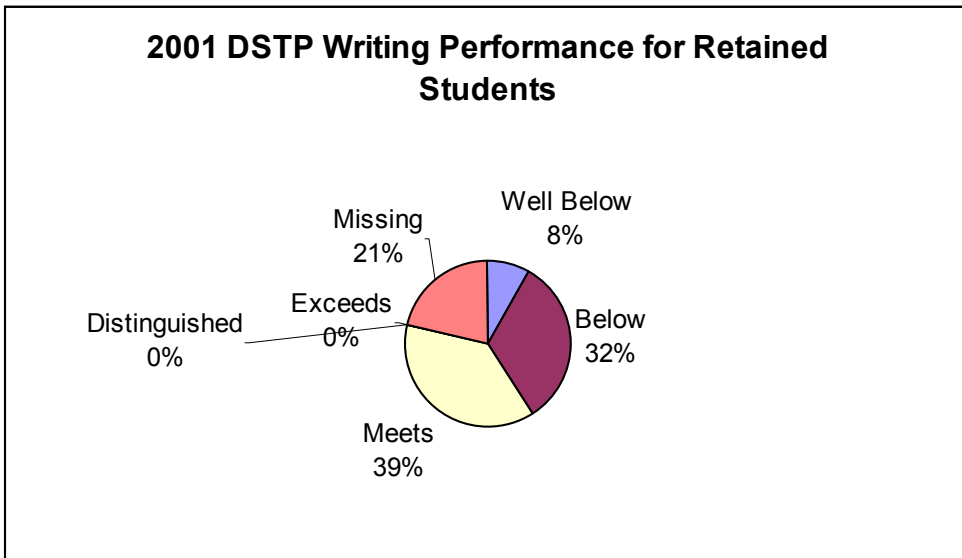
Graph 84



Graph 85



Graph 86



An examination of the data revealed that larger percentages of retained students met the reading standard in 2001 but that one-third of the students retained in grade 8 were performing below the standard in 2001. Larger percentages of retained students were also meeting the math standard in 2001 but a majority of the retained students were performing below the standard in 2001. With respect to the writing portion of the DSTP, larger percentages of retained students were meeting the standard in 2001 but more than one-third were performing below the standard in 2001.

COMPLETION RATES

In addition to examining special education placement rates, student achievement, student behavior, and retention rates within the Delaware student accountability system, the current study was also intended to examine drop-out rates within the system. More specifically, the questions to be addressed with respect to drop-outs included:

- **What are the demographic characteristics of students who dropout?**
- **What is the relationship of dropping-out to performance on the DSTP?**

These questions can only be addressed based on the secondary cohort of students because, by law, elementary students are not permitted to drop-out of school. Even within the secondary cohort there were not expected to be many official drop-outs reported for 1999-00 or 2000-01 because the drop-out rates for 7th and 8th grades have been historically low in Delaware.¹⁰

DEMOGRAPHIC CHARACTERISTICS OF DROP-OUTS

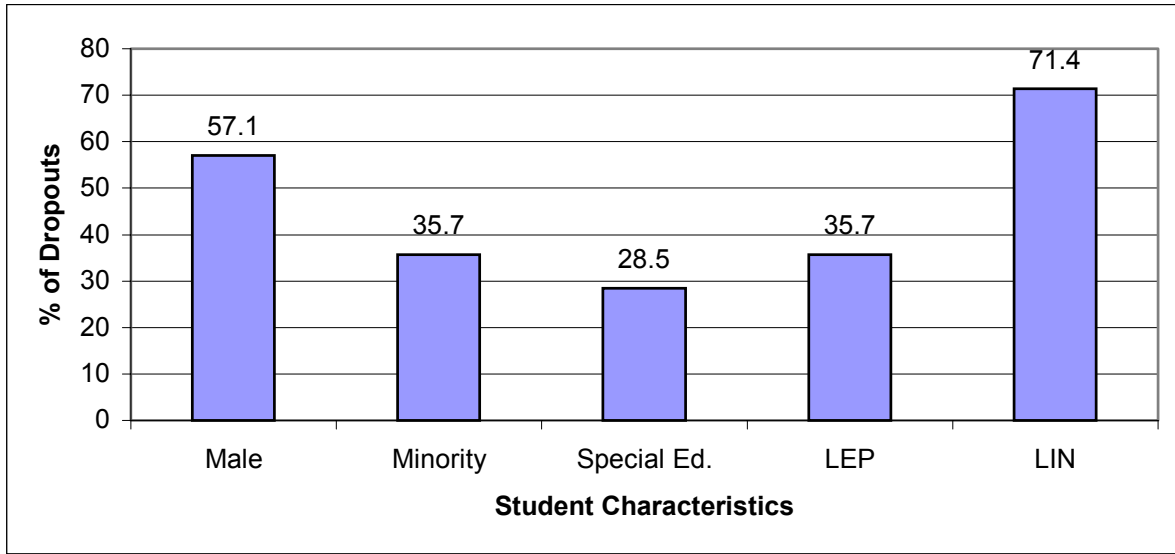
In 1999-00 there were 4 official drop-outs reported for the secondary cohort. The students who dropped out had the following characteristics:

- Caucasian, low income male who had previously been retained;
- Caucasian, low income female who had previously been retained;
- Hispanic, special education male who was in the appropriate grade at the time of dropping out; and,
- Caucasian male who was in the appropriate grade at the time of dropping out.

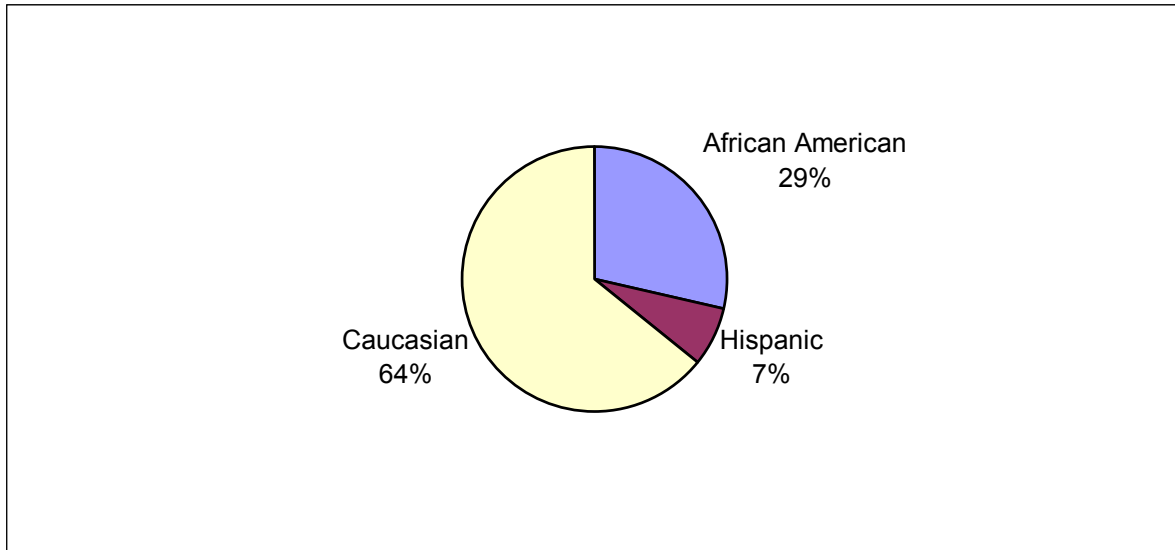
Given that there were so few drop-outs during the 1999-00 school year no real statement of the relationship between demographic characteristics and dropping-out could be made.

During the 2000-01, 14 students were identified as officially dropping out. The demographic characteristics of these students according to their most up-to-date enrollment information was as follows:

Graph 87: Student Demographic information for Dropouts in 2000-01



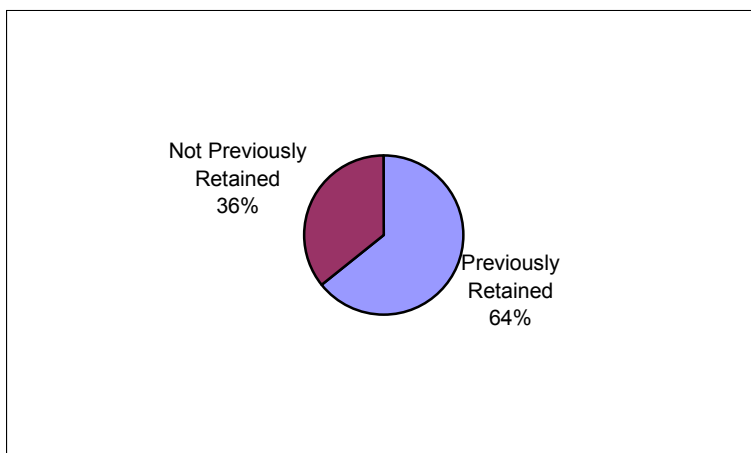
Graph 88: Ethnicity of 8th Grade Drop-outs



Although more students dropped out in 2000-01 the relatively small number prevents a statistical examination of the relationship between student demographics and dropping out. A cursory examination of the demographic information of the 2000-01 drop-outs suggests that minority students may not be over-represented among drop-outs but low income students may.

An examination of the promotion history of the 14 students who dropped-out in 2000-01 revealed that nearly two-thirds of these students had been previously retained during the course of this study (See Graph 89). This finding lends support to the pre-existing literature on retention which indicates that students who are retained are more likely to drop-out.

Graph 89: Promotion History of Eventual Drop-outs

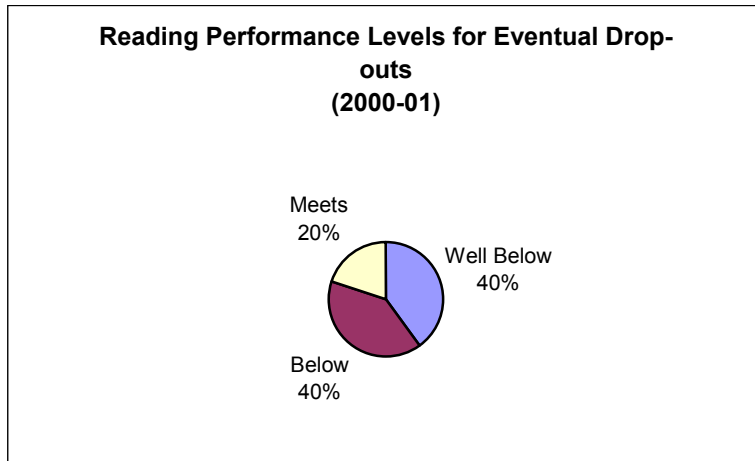


DROPPING OUT AND DSTP PERFORMANCE

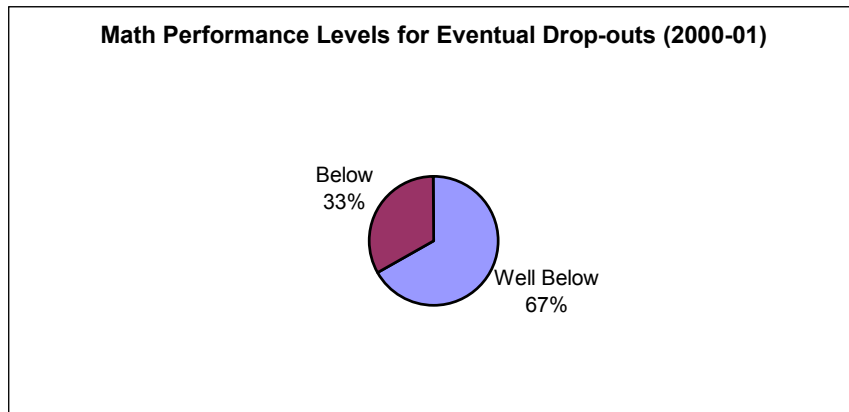
The relationship of dropping out to performance on the DSTP could only be examined in a descriptive manner given the small number of official drop-outs in the secondary cohort. Of the 4 students who dropped-out during the 1999-00 school year, only one student had progressed through the system in time to take the DSTP in the Spring of 2000 and had a score that could be aggregated. This student scored well below the standard in all three content areas.

Of the 14 students who dropped out during the 2000-01 school year, less than half of these had DSTP scores that could be examined. The DSTP performance for these few students were as follows:

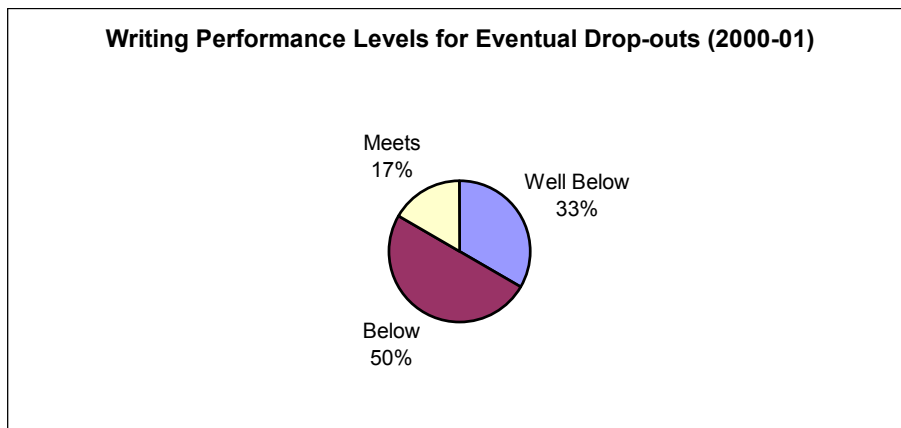
Graph 90



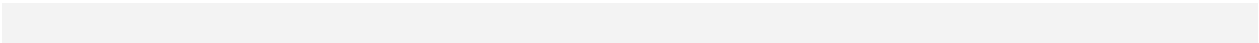
Graph 91



Graph 92



Keeping the small sample size in consideration, a review of the DSTP performance for drop-outs during the 2000-01 school year revealed that most of these students were performing below the standard in all three content areas.



POLICY CONSIDERATIONS

Issue 1: Special Education Placement Rates

At this stage in the monitoring process of Delaware's student accountability plan there do not appear to be increases special education rates prior to testing. In fact, the opposite has occurred but in a systematic manner. The largest percentages of re-classified students have been those students whose classification changed from special education to regular education at the time of test. These reclassifications generally involve speech impaired students who were re-classified as regular education for testing purposes.

While tracking special education placement rates over time is a good starting point for monitoring the accountability system, a further investigation of the testing conditions for special education students should be conducted. More information concerning student participation rates and the use of accommodations that do not permit aggregation should be provided.

Policy Considerations:

- What safeguards are in place to prevent the overuse of accommodations that prohibit aggregations?
- On the other hand, in light of the school accountability system and the assigning of performance levels of 0 to students who test with accommodations that do not allow for aggregation, what safeguards are in place to ensure that students with special needs will receive the appropriate accommodations?

Issue 2: Student Achievement

Analyses based on cross-sectional examinations of student performance on the DSTP revealed that improvements occurred for most student groups at the elementary level in reading and mathematics but that fewer student groups showed improvement at the secondary level. An examination of the percentage of students at each of the performance levels revealed that race and income gaps in achievement continue to persist but appear to be narrowing in the elementary grades. The gaps in 10th grade mathematics grew slightly for most student groups examined. In addition, the percentage of students at or above the standard was generally higher in the elementary grades and smaller in the secondary grades and was found to be higher for reading than math.

Policy considerations:

- What lessons are there to be learned from the Title I programs given the improvements in DSTP performance evidenced by Title I students in grades 3, 5, and 8?

- How does the state plan to respond to the disparate impact of the standards-led reform on particular groups of Delaware students (i.e. African Americans, Hispanics, Low Income students, 10th Graders)?

Issue 3: Behavioral Effects

The results of this year's report concerning out-of-school suspension echo those reported last year. Analyses of demographic characteristics of suspended students again revealed that the odds of being suspended were greater for male, low income, minority students. In addition, the results indicated that the rate of suspensions were on the rise within the elementary cohort.

Out-of-school suspension is considered to be one of the least effective strategies for improving student behavior but is widely used. In high-stakes settings, removing a student from the school is particularly problematic in that it may limit a student's opportunity to learn. With respect to DSTP performance, a majority of the students receiving suspensions during the 1999-00 school year performed below state content standards in reading and mathematics.

Policy Considerations:

- Given the research base indicating the ineffectiveness of out-of-school suspension, what other means of dealing with misconduct should be promoted that would have fewer negative effects on student performance?
- To what extent are suspended students provided with adequate opportunities to learn? What measures are currently taken to ensure that suspended students do not fall further behind in school?
- What attempts have been made to monitor the rate of out-of-school suspensions during the DSTP testing period?

Issue 4: Retention

Like out-of-school suspension, the practice of retaining students in grade has been found to be an ineffective strategy for assisting under-performing students. The results of this year's investigation revealed that, as with out-of-school suspension, low income, minority, and male students were more likely to be retained. An examination of the DSTP performance levels of retained students revealed increases in the percentages of students meeting the standard following retention but that larger percentages of the retained students were below the standard in math and more than one-third of the students were below the standard in reading in writing.

Policy Considerations:

- What role do the Individualized Student Improvement Plans (IIPs) play in assisting under-performing students? How does the state plan to evaluate the effectiveness of the IIPs?
- Beyond summer school and IIPs, what other strategies are currently being employed to assist under-performing students and to what degree are they being successful?

Issue 5: Completion Rates

To this point in the monitoring process the number of students who have officially dropped out of school is relatively small. The data available on these drop-outs suggest that low income students may be over-represented among drop-outs and that students with a history or retention may be more likely to drop-out. Preliminary analyses of the DSTP data for the drop-outs indicated that most of the eventual drop-outs were performing below the standard in reading, mathematics and writing.

Policy Consideration:

- Given the relationship between retention and increased risks of dropping-out, what attempts are being made to identify likely drop-outs and provide these students with assistance?

REFERENCES

1. National Research Council. (1999). *High Stakes: Testing for Tracking, Promotion, and Graduation*. Washington, DC: National Academy Press.
2. Fuhrman, S. (1999). The New Accountability. (CPRE Policy Brief No. RB-27). Philadelphia, PA: University of Pennsylvania, Consortium for Policy research in Education. Allington, R., & McGill, Frazen, A. (1992). Unintended effects of educational reform in New York. *Educational Policy*, 6, 397-414.
3. Personal communication with Robin Taylor (Fall 2001) and George Smith (Spring 2001).
4. Noble, A.J., & Banicky, L. (2000). *Synchronizing the Accountability Clocks: A Policy Study of Delaware's Student Accountability Plan*. <http://www.rdc.udel.edu>
5. Bailey, K., & Pilhofer, A. (July 23, 2001). Change holds down third-grade scores. Delaware News Journal.
6. Moles, O. (1989). *Strategies to Reduce Student Misbehavior*. Washington, DC: Office of Educational Research and Improvement. Slee, R. (1986). Integration: The Disruptive Student and Suspension. *The Urban Review*, 18, 87-103.
7. Ambrose, M., & Gibson, M. (1995). Does suspension work? *NEA Today*, 13(7), 39.
8. Garibaldi, A., Blanchard, L., & Brooks, S. (1998). Conflict resolution training, teacher effectiveness, and student suspension: The impact of a health and safety initiative in the New Orleans Public Schools. *Journal of Negro Education*, 65(4), 408-413. Skiba, R., Peterson, R., & Williams, T. (1997). Office referrals and suspensions: Disciplinary interventions in middle schools. *Education and treatment of children*, 20(3), 295-315.
9. U.S. Department of Education. (1999). *Taking Responsibility for ending social promotion*. Washington, DC: Author. Dill, S.V. (1993). Closing the Gap: Acceleration vs. remediation and the impact of retention in the grade on student achievement. (ERIC Document Reproduction Services No. 364 938).
10. Annual Dropout Statistics Reports for Delaware. Available at http://www.doe.state.de.us/reports/School_Reports.htm

APPENDIX A

Table 1: Cohort Profiles

	Elementary (3 rd Graders in 1997-98) (N=8404)	Secondary (6 th Graders in 1997-989) (N=8732)
Male	53%	52%
Female	47%	48%
American Indian	.3%	.3%
African American	30.5%	30.6%
Asian	1.7%	1.7%
Hispanic	4.9%	4.6%
Caucasian	62.6%	62.9%
Special Education	17%	15%
Low Income	44%	42%
Title I	12%	4%
LEP	2%	1.3%

APPENDIX B

TECHNICAL DETAILS CONCERNING THE DSTP

In order to address the student achievement issue, a few technical details concerning the DSTP should be addressed. The DSTP is composed of multiple choice, short answer, and extended response items. Results are reported out in the form of national percentile ranks, standards based scores, and performance levels.

PERCENTILE RANKS

The national percentile ranking are based on abbreviated versions of the reading comprehension and the mathematical problem solving subsets of the Stanford Achievement Test series, 9th Edition (SAT9). The SAT9 is a norm-referenced test published by Harcourt Brace Educational Measurement. Percentile ranks cannot be manipulated mathematically because there are not equal intervals between percentile ranks. For example, the difference between a percentile rank of 5 and 10 is not the same as the difference in achievement as the difference between a percentile rank of 50 and 55. This point is worth noting not only for the analyses that follow but also for the purpose of avoiding incorrect conclusions based on cursory examinations of data.

In order to be used in statistical analyses the national percentile rankings must be converted to another metric, in this case normal curve equivalents(NCEs). NCEs provide an equal-interval scale and therefore allow for mathematical manipulation. NCEs can range from 1 to 99.

A WORD OF CAUTION ABOUT STATISTICAL INTERPRETATIONS

The goal of many statistical analyses is to show that there is some difference between sets of observations, and that the difference is due to something other than chance factors. For example, when examining the elementary cohort's average SAT9 math score from 1998 (Mean=53.88) to their average SAT9 math score in 2000 (Mean=57.35) there is an increase of 3.43. Finding such a difference does not necessarily mean that it is a meaningful difference. This difference may simply reflect the amount of variability in the data. Statistical analyses are set up in such a way as to compare the difference found to a measure of how much variability can be expected simply due to chance factors. The extent to which the difference is over and above the amount of variability expected simply due to chance determines whether or not a result is statistically significant. Therefore a statistically significant result simply means that an outcome, in this example a difference of 3.43, is unlikely to be due to chance factors and instead may represent an actual improvement in scores.

Recently many researchers have argued that significance tests can be misleading because with very large sample sizes, even the smallest difference between two sets of observations can result in a significant finding. Therefore testing for statistical significance is often viewed as the first step in data analysis with the second step focused on the size of the “effect”.¹

To use an analogy, testing for statistical significance is like using a magnifying glass to locate an object. The size of the sample determines the “magnification” of the lens in the magnifying glass such that larger samples result in even the smallest object appearing quite large. Effect size can be thought of as a ruler that researchers use to measure the size of their findings. In the case of the magnifying glass analogy, we may locate an object that appears to be quite large, but when the ruler (i.e. effect size) is placed next to it under the magnifying glass the size of the object is placed in a more meaningful context. Conventional rules of thumb indicate that an effect size of .2 is small, an effect size of .5 is medium, and an effect size of .8 is large.²

¹ See Aron, A., & Aron, E. (1994). *Statistics for Psychology*, pp. 239-240), Upper Saddle River, NJ: Prentice Hall.

² Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

APPENDIX C

